

(19) 世界知的所有権機関
国際事務局(43) 国際公開日
2002年2月14日 (14.02.2002)

PCT

(10) 国際公開番号
WO 02/13183 A1

(51) 国際特許分類: G10L 13/00, 19/04 // 101:12

(21) 国際出願番号: PCT/JP01/06708

(22) 国際出願日: 2001年8月3日 (03.08.2001)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
 特願2000-241062 2000年8月9日 (09.08.2000) JP
 特願2000-251969 2000年8月23日 (23.08.2000) JP
 特願2000-346675 2000年11月14日 (14.11.2000) JP

(71) 出願人(米国を除く全ての指定国について): ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒141-0001 東京都品川区北品川6丁目7番35号 Tokyo (JP).

(72) 発明者; および
 (75) 発明者/出願人(米国についてのみ): 近藤哲二郎

(KONDO, Tetsujiro) [JP/JP]. 渡辺 勉 (WATANABE, Tsutomu) [JP/JP]. 服部 正明 (HATTORI, Masaaki) [JP/JP]. 木村裕人 (KIMURA, Hiroto) [JP/JP]. 藤森泰弘 (FUJIMORI, Yasuhiro) [JP/JP]; 〒141-0001 東京都品川区北品川6丁目7番35号 ソニー株式会社内 Tokyo (JP).

(74) 代理人: 小池 晃, 外 (KOIKE, Akira et al); 〒105-0001 東京都港区虎ノ門二丁目6番4号 第11森ビル Tokyo (JP).

(81) 指定国(国内): KR, NO, US.

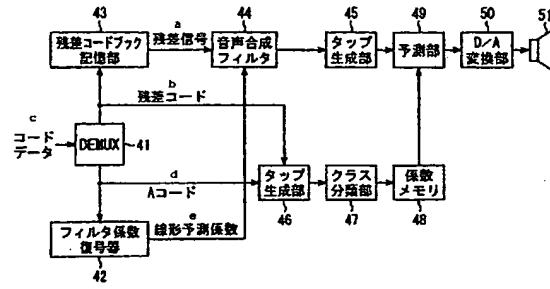
(84) 指定国(広域): ヨーロッパ特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

添付公開書類:
 — 国際調査報告書

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイドスノート」を参照。

(54) Title: VOICE DATA PROCESSING DEVICE AND PROCESSING METHOD

(54) 発明の名称: 音声データの処理装置及び処理方法



43...RESIDUAL CODE BOOK STORAGE SECTION	44...SPEECH SYNTHESIZATION FILTER
a...RESIDUAL SIGNAL	45...TAP CREATING SECTION
b...RESIDUAL CODE	46...TAP CREATING SECTION
c...CODE DATA	47...CATEGORIZING SECTION
d...CODE A	48...COEFFICIENT MEMORY
e...FILTER COEFFICIENT DECODER	
f...LINEAR PREDICTION COEFFICIENT	

WO 02/13183 A1

(57) Abstract: A voice processing device for determining a prediction value of a voice of high sound quality by extracting a prediction tap for predicting a prediction value of a voice of high sound quality from a synthesized voice produced by imparting a linear prediction coefficient determined from a predetermined code and a residual signal to a speech synthesizing filter and by performing a predetermined calculation by using the prediction tap and a predetermined tap coefficient, comprising a prediction tap extracting section (45) for extracting from a synthesized voice a prediction tap used for predicting a voice of interest of high sound quality for which a prediction value is to be determined, a class tap extracting section (46) for extracting a class tap used for categorizing the voice of interest into one of class from a code, a categorizing section (47) for categorizing the voice of interest into a classes on the basis of the class tap, a tap creating section for acquiring a tap coefficient corresponding to the class of the voice of interest from tap coefficients of the respective classes determined by learning, and a prediction section (49) for determining a prediction value of the voice of interest by using the prediction tap and the tap coefficient corresponding to the class of the voice of interest.

(続葉有)



(57) 要約:

本発明は、所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音から、その音質を向上させた高音質の音声の予測値を予測するための予測タップを抽出し、その予測タップと所定のタップ係数を用いて、所定の予測演算を行うことにより、高音質の音声の予測値を求める音声処理装置であり、予測値を求めようとしている高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップを、合成音から抽出する予測タップ抽出部（45）と、注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、コードから抽出するクラスタップ抽出部（46）と、クラスタップに基づいて、注目音声のクラスを求めるクラス分類を行うクラス分類部（47）と、学習を行うことにより求められたクラスごとのタップ係数の中から注目音声のクラスに対応するタップ係数を取得するタップ生成部と、予測タップと注目音声のクラスに対応するタップ係数を用いて注目音声の予測値を求める予測部（49）とを備える。

明細書

音声データの処理装置及び処理方法

技術分野

本発明は、 本発明は、データ処理装置及びデータ処理方法、学習装置及び学習方法、並びに記録媒体に関し、特に、例えば、C E L P (Code Excited Liner Prediction coding) 方式で符号化された音声を、高音質の音声に復号することができるようとするデータ処理装置及びデータ処理方法、学習装置及び学習方法、並びに記録媒体に関する。

背景技術

まず、従来用いられている携帯電話機の一例を図1及び図2を参照して説明する。

この携帯電話機では、音声を、C E L P方式により所定のコードに符号化して送信する送信処理と、他の携帯電話機から送信されてくるコードを受信して、音声に復号する受信処理とが行われるようになっており、図1は、送信処理を行う送信部を、図2は、受信処理を行う受信部を、それぞれ示している。

図1に示した送信部では、ユーザが発話した音声が、マイクロフォン1に入力され、そこで、電気信号としての音声信号に変換され、A/D (Analog/Digital) 変換部2に供給される。A/D 変換部2は、マイクロフォン1からのアナログの音声信号を、例えば、8 kHz等のサンプリング周波数でサンプリングすることにより、デジタルの音声信号にA/D 変換し、さらに、所定のビット数で量子化を行って、演算器3とLPC (Liner Prediction Coefficient) 分析部4に供給する。

LPC 分析部4は、A/D 変換部2からの音声信号を、例えば、160サンプル分の長さのフレームごとにLPC分析し、P次の線形予測係数 $\alpha_1, \alpha_2, \dots$

・、 α_p を求める。そして、LPC分析部4は、このP次の線形予測係数 α_p ($p = 1, 2, \dots, P$) を要素とするベクトルを、音声の特徴ベクトルとして、ベクトル量子化部5に供給する。

ベクトル量子化部5は、線形予測係数を要素とするコードベクトルとコードとを対応付けたコードブックを記憶しており、そのコードブックに基づいて、LPC分析部4からの特徴ベクトル α をベクトル量子化し、そのベクトル量子化の結果得られるコード（以下、適宜、Aコード(A_code)という）を、コード決定部15に供給する。

さらに、ベクトル量子化部5は、Aコードに対応するコードベクトル α' を構成する要素となっている線形予測係数 $\alpha_1', \alpha_2', \dots, \alpha_P'$ を、音声合成フィルタ6に供給する。

音声合成フィルタ6は、例えば、IIR (Infinite Impulse Response)型のデジタルフィルタで、ベクトル量子化部5からの線形予測係数 α_p' ($p = 1, 2, \dots, P$) をIIRフィルタのタップ係数とするとともに、演算器14から供給される残差信号 e を入力信号として、音声合成を行う。

即ち、LPC分析部4で行われるLPC分析は、現在時刻nの音声信号（のサンプル値） s_n 、及びこれに隣接する過去のP個のサンプル値 $s_{n-1}, s_{n-2}, \dots, s_{n-P}$ に、式(1)

$$s_n + \alpha_1 s_{n-1} + \alpha_2 s_{n-2} + \dots + \alpha_P s_{n-P} = e_n \quad \dots \quad (1)$$

で示す線形1次結合が成立すると仮定し、現在時刻nのサンプル値 s_n の予測値（線形予測値） s_n' を、過去のP個の標本値 $s_{n-1}, s_{n-2}, \dots, s_{n-P}$ を用いて、式(2)

$$s_n' = -(\alpha_1 s_{n-1} + \alpha_2 s_{n-2} + \dots + \alpha_P s_{n-P}) \quad \dots \quad (2)$$

によって線形予測したときに、実際のサンプル値 s_n と線形予測値 s_n' との間の自乗誤差を最小にする線形予測係数 α_p を求めるものである。

ここで、式(1)において、 $\{e_n\} (\dots, e_{n-1}, e_n, e_{n+1}, \dots)$ は、平均値が0で、分散が所定値 σ^2 の互いに無相関な確率変数である。

式(1)から、サンプル値 s_n は、式(3)

$$s_n = e_n - (\alpha_1 s_{n-1} + \alpha_2 s_{n-2} + \dots + \alpha_P s_{n-P}) \quad \dots \quad (3)$$

で表すことができ、これを、Z変換すると、次式が成立する。

$$S = E / (1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots + \alpha_p z^{-p}) \quad \dots \quad (4)$$

但し、式(4)において、SとEは、式(3)における s_n と e_n のZ変換を、それぞれ表す。

ここで、式(1)及び(2)から、 e_n は、式(5)

$$e_n = s_n - s_n' \quad \dots \quad (5)$$

で表すことができ、実際のサンプル値 s_n と線形予測値 s_n' との間の残差信号と呼ばれる。

従って、式(4)から、線形予測係数 α_p をIIRフィルタのタップ係数とともに、残差信号 e_n をIIRフィルタの入力信号とすることにより、音声信号 s_n を求めることができる。

音声合成フィルタ6は、上述したように、ベクトル量子化部5からの線形予測係数 α_p' をタップ係数とともに、演算器14から供給される残差信号 e を入力信号として、式(4)を演算し、音声信号(合成音信号) s_s を求める。

なお、音声合成フィルタ6では、LPC分析部4によるLPC分析の結果得られる線形予測係数 α_p ではなく、そのベクトル量子化の結果得られるコードに対応するコードベクトルとしての線形予測係数 α_p' が用いられるため、音声合成フィルタ6が出力する合成音信号は、A/D変換部2が出力する音声信号とは、基本的に同一にはならない。

音声合成フィルタ6が出力する合成音信号 s_s は、演算器3に供給される。演算器3は、音声合成フィルタ6からの合成音信号 s_s から、A/D変換部2が出力する音声信号 s を減算し、その減算値を、自乗誤差演算部7に供給する。自乗誤差演算部7は、演算器3からの減算値の自乗和(第kフレームのサンプル値についての自乗和)を演算し、その結果得られる自乗誤差を、自乗誤差最小判定部8に供給する。

自乗誤差最小判定部8は、自乗誤差演算部7が出力する自乗誤差に対応付けて、ラグを表すコードとしてのLコード(L_code)、ゲインを表すコードとしてのGコード(G_code)、及び符号語を表すコードとしてのIコード(I_code)を記憶しており、自乗誤差演算部7が出力する自乗誤差に対応するLコード、Gコード、及び

及びLコードを出力する。Lコードは、適応コードブック記憶部9に、Gコードは、ゲイン復号器10に、Iコードは、励起コードブック記憶部11に、それぞれ供給される。さらに、Lコード、Gコード、及びIコードは、コード決定部15にも供給される。

適応コードブック記憶部9は、例えば7ビットのLコードと、所定の遅延時間(ラグ)とを対応付けた適応コードブックを記憶しており、演算器14から供給される残差信号eを、自乗誤差最小判定部8から供給されるLコードに対応付けられた遅延時間だけ遅延して、演算器12に出力する。

ここで、適応コードブック記憶部9は、残差信号eを、Lコードに対応する時間だけ遅延して出力することから、その出力信号は、その遅延時間を周期とする周期信号に近い信号となる。この信号は、線形予測係数を用いた音声合成において、主として、有声音の合成音を生成するための駆動信号となる。

ゲイン復号器10は、Gコードと、所定のゲイン β 及び γ とを対応付けたテーブルを記憶しており、自乗誤差最小判定部8から供給されるGコードに対応付けられたゲイン β 及び γ を出力する。ゲイン β と γ は、演算器12と13に、それぞれ供給される。

励起コードブック記憶部11は、例えば9ビットのIコードと、所定の励起信号とを対応付けた励起コードブックを記憶しており、自乗誤差最小判定部8から供給されるIコードに対応付けられた励起信号を、演算器13に出力する。

ここで、励起コードブックに記憶されている励起信号は、例えば、ホワイトノイズ等に近い信号であり、線形予測係数を用いた音声合成において、主として、無声音の合成音を生成するための駆動信号となる。

演算器12は、適応コードブック記憶部9の出力信号と、ゲイン復号器10が出力するゲイン β とを乗算し、その乗算値1を、演算器14に供給する。演算器13は、励起コードブック記憶部11の出力信号と、ゲイン復号器10が出力するゲイン γ とを乗算し、その乗算値nを、演算器14に供給する。演算器14は、演算器12からの乗算値1と、演算器13からの乗算値nとを加算し、その加算値を、残差信号eとして、音声合成フィルタ6に供給する。

音声合成フィルタ6では、以上のようにして、演算器14から供給される残差

信号 e を入力信号が、ベクトル量子化部5から供給される線形予測係数 α_p' をタップ係数とするIIRフィルタでフィルタリングされ、その結果得られる合成音信号が、演算器3に供給される。そして、演算器3及び自乗誤差演算部7において、上述の場合と同様の処理が行われ、その結果得られる自乗誤差が、自乗誤差最小判定部8に供給される。

自乗誤差最小判定部8は、自乗誤差演算部7からの自乗誤差が最小（極小）になったかどうかを判定する。そして、自乗誤差最小判定部8は、自乗誤差が最小になっていないと判定した場合、上述のように、その自乗誤差に対応するLコード、Gコード、及びIコードを出力し、以下、同様の処理が繰り返される。

一方、自乗誤差最小判定部8は、自乗誤差が最小になったと判定した場合、確定信号を、コード決定部15に出力する。コード決定部15は、ベクトル量子化部5から供給されるAコードをラッチするとともに、自乗誤差最小判定部8から供給されるLコード、Gコード、及びIコードを順次ラッチするようになっており、自乗誤差最小判定部8から確定信号を受信すると、そのときラッチしているAコード、Lコード、Gコード、及びIコードを、チャネルエンコーダ16に供給する。チャネルエンコーダ16は、コード決定部15からのAコード、Lコード、Gコード、及びIコードを多重化し、コードデータとして出力する。このコードデータは、伝送路を介して送信される。

以下では、説明を簡単にするため、Aコード、Lコード、Gコード、及びIコードは、フレームごとに求められるものとする。但し、例えば、1フレームを、4つのサブフレームに分割し、Lコード、Gコード、及びIコードは、サブフレームごとに求めるようにすること等が可能である。

ここで、図1（後述する図2、図11、及び図12においても同様）では、各変数に、[k]が付され、配列変数とされている。このkは、フレーム数を表すが、明細書中では、その記述は、適宜省略する。

以上のようにして、他の携帯電話機の送信部から送信されてくるコードデータは、図2に示した受信部のチャネルデコーダ21で受信される。チャネルデコーダ21は、コードデータから、Lコード、Gコード、Iコード、Aコードを分離し、それぞれを、適応コードブック記憶部22、ゲイン復号器23、励起コード

ブック記憶部 24、フィルタ係数復号器 25 に供給する。

適応コードブック記憶部 22、ゲイン復号器 23、励起コードブック記憶部 24、演算器 26 乃至 28 は、図 1 の適応コードブック記憶部 9、ゲイン復号器 10、励起コードブック記憶部 11、演算器 12 乃至 14 とそれ同様に構成されるもので、図 1 で説明した場合と同様の処理が行われることにより、L コード、G コード、及び I コードが、残差信号 e に復号される。この残差信号 e は、音声合成フィルタ 29 に対して、入力信号として与えられる。

フィルタ係数復号器 25 は、図 1 のベクトル量子化部 5 が記憶しているのと同一のコードブックを記憶しており、A コードを、線形予測係数 α_{ν}' に復号し、音声合成フィルタ 29 に供給する。

音声合成フィルタ 29 は、図 1 の音声合成フィルタ 6 と同様に構成されており、フィルタ係数復号器 25 からの線形予測係数 α_{ν}' をタップ係数とするとともに、演算器 28 から供給される残差信号 e を入力信号として、式 (4) を演算し、これにより、図 1 の自乗誤差最小判定部 8 において自乗誤差が最小と判定されたときの合成音信号を生成する。この合成音信号は、D/A (Digital/Analog) 変換部 30 に供給される。D/A 変換部 30 は、音声合成フィルタ 29 からの合成音信号を、デジタル信号からアナログ信号に D/A 変換し、スピーカ 31 に供給して出力させる。

以上のように、携帯電話機の送信部では、受信部の音声合成フィルタ 29 に与えられるフィルタデータとしての残差信号と線形予測係数がコード化されて送信されてくるため、受信部では、そのコードが、残差信号と線形予測係数に復号される。この復号された残差信号や線形予測係数（以下、適宜、それを、復号残差信号又は復号線形予測係数という）には、量子化誤差等の誤差が含まれるため、音声を L P C 分析して得られる残差信号と線形予測係数には一致しない。このため、受信部の音声合成フィルタ 29 が output する合成音信号は、歪みを有する、音質の劣化したものとなる。

発明の開示

本発明は、上述したような実情に鑑み提案されるものであって、本発明の目的とするところは、高音質の合成音を得ることができる音声データの処理装置及びデータ処理方法、更には、これらデータ処理装置及び方法を用いた学習装置及び学習方法を提供することにある。

上述のような目的を達成するために提案される本発明に係る音声処理装置は、予測値を求めようとしている高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップを、合成音から抽出する予測タップ抽出部と、注目音声を幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップをコードから抽出するクラスタップ抽出部と、クラスタップに基づいて注目音声のクラスを求めるクラス分類を行うクラス分類部と、学習を行うことにより求められたクラスごとのタップ係数の中から注目音声のクラスに対応するタップ係数を取得する取得部と、予測タップと、注目音声のクラスに対応するタップ係数を用いて、注目音声の予測値を求める予測部とを備え、予測値を求めようとしている高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップを合成音から抽出し、注目音声を幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップをコードから抽出し、クラスタップに基づいて注目音声のクラスを求めるクラス分類し、学習を行うことにより求められたクラスごとのタップ係数の中から注目音声のクラスに対応するタップ係数を取得し、予測タップと注目音声のクラスに対応するタップ係数を用いて注目音声の予測値を求める。

本発明に係る学習装置は、予測値を求めようとしている高音質の音声を注目音声として、その注目音声を幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップをコードから抽出するクラスタップ抽出部と、クラスタップに基づいて注目音声のクラスを求めるクラス分類を行うクラス分類部と、タップ係数及び合成音を用いて予測演算を行うことにより得られる高音質の音声の予測値の予測誤差が統計的に最小になるように学習を行い、クラスごとのタップ係数を求める学習手段とを備え、予測値を求めようとしている高音質の音声を注目音声として、その注目音声を幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、コードから抽出し、クラスタップに基づいて注目音

声のクラスを求めるクラス分類し、タップ係数及び合成音を用いて予測演算を行うことにより得られる高音質の音声の予測値の予測誤差が統計的に最小になるように学習を行い、クラスごとのタップ係数を求める。

また、本発明に係るデータ処理装置は、コードを復号し、復号フィルタデータを出力するコード復号部と、学習を行うことにより求められた所定のタップ係数を取得する取得部と、タップ係数及び復号フィルタデータを用いて所定の予測演算を行うことによりフィルタデータの予測値を求め、音声合成フィルタに供給する予測部とを備え、コードを復号し、復号フィルタデータを出力し、学習を行うことにより求められた所定のタップ係数を取得し、タップ係数及び復号フィルタデータを用いて所定の予測演算を行うことによりフィルタデータの予測値を求め音声合成フィルタに供給する。

さらに、本発明に係る学習装置は、フィルタデータに対応するコードを復号し、復号フィルタデータを出力するコード復号部と、タップ係数および復号フィルタデータを用いて予測演算を行うことにより得られるフィルタデータの予測値の予測誤差が統計的に最小になるように学習を行いタップ係数を求める学習手段とを備え、フィルタデータに対応するコードを復号し、復号フィルタデータを出力するコード復号ステップと、タップ係数および復号フィルタデータを用いて予測演算を行うことにより得られるフィルタデータの予測値の予測誤差が統計的に最小になるように学習を行う。

本発明に係る音声処理装置は、予測値を求めようとしている高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップを、合成音と、コードまたはコードから得られる情報とから抽出する予測タップ抽出部と、注目音声を幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、合成音とコード又はコードから得られる情報とから抽出するクラスタップ抽出部と、クラスタップに基づいて注目音声のクラスを求めるクラス分類を行うクラス分類部と、学習を行うことにより求められたクラスごとのタップ係数の中から注目音声のクラスに対応するタップ係数を取得する取得部と、予測タップと注目音声のクラスに対応するタップ係数を用いて注目音声の予測値を求める予測部とを備え、予測値を求めようとしている高音質の音声を注目音声として、その注

目音声を予測するのに用いる予測タップを、合成音とコード又はコードから得られる情報とから抽出し、注目音声を幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、合成音と、コード又はコードから得られる情報とから抽出し、クラスタップに基づいて注目音声のクラスを求めるクラス分類を行、学習を行うことにより求められたクラスごとのタップ係数の中から注目音声のクラスに対応するタップ係数を取得し、予測タップと注目音声のクラスに対応するタップ係数を用いて、注目音声の予測値を求める。

また、本発明に係る学習装置は、予測値を求めようとしている高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップを、合成音と、コード又はコードから得られる情報とから抽出する予測タップ抽出部と、注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、合成音と、コードまたはコードから得られる情報とから抽出するクラスタップ抽出部と、クラスタップに基づいて、注目音声のクラスを求めるクラス分類を行うクラス分類部と、タップ係数及び予測タップを用いて予測演算を行うことにより得られる高音質の音声の予測値の予測誤差が統計的に最小になるように学習を行い、クラスごとのタップ係数を求める学習手段とを備え、予測値を求めようとしている高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップを、合成音と、コード又はコードから得られる情報とから抽出し、注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、合成音と、コードまたはコードから得られる情報とから抽出し、クラスタップに基づいて注目音声のクラスを求めるクラス分類を行い、タップ係数及び予測タップを用いて予測演算を行うことにより得られる高音質の音声の予測値の予測誤差が統計的に最小になるように学習を行い、クラスごとのタップ係数を求める。

本発明の更に他の目的、本発明によって得られる具体的な利点は、以下に説明される実施例の説明から一層明らかにされるであろう。

図面の簡単な説明

図1は、従来の携帯電話機を構成する送信部の一例を示すブロック図であり、図2は、受信部の一例を示すブロック図である。

図3は、本発明を適用した音声合成装置を示すブロック図であり、図4は、音声合成装置を構成する音声合成フィルタを示すブロック図である。

図5は、図3に示す音声合成装置の処理を説明するフローチャートである。

図6は、本発明を適用した学習装置を示すブロック図である。

図7は、本発明に学習装置を構成する予測フィルタを示すブロック図である。

図8は、図6に示す学習装置の処理を説明するフローチャートである。

図9は、本発明を適用した伝送システムを示すブロック図である。

図10は、本発明が適用された携帯電話機を示すブロック図である。

図11は、携帯電話機を構成する受信部を示すブロック図である。

図12は、本発明を適用した学習装置の他を示すブロック図である。

図13は、本発明を適用したコンピュータの一構成例を示すブロック図である。

図14は、本発明を適用した音声合成装置の他の例を示すブロック図であり、

図15は、音声合成装置を構成する音声合成フィルタを示すブロック図である。

図16は、図14に示す音声合成装置の処理を説明するフローチャートである。

図17は、本発明を適用した学習装置の他の例を示すブロック図である。

図18は、本発明に学習装置を構成する予測フィルタを示すブロック図である。

図19は、図17に示す学習装置の処理を説明するフローチャートである。

図20は、本発明を適用した伝送システムを示すブロック図である。

図21は、本発明が適用された携帯電話機を示すブロック図である。

図22は、携帯電話機を構成する受信部を示すブロック図である。

図23は、本発明を適用した学習装置の他を示すブロック図である。

図24は、本発明を適用した音声合成装置のさらに他の例を示すブロック図であり、図25は、音声合成装置を構成する音声合成フィルタを示すブロック図である。

図26は、図24に示す音声合成装置の処理を説明するフローチャートである。

図27は、本発明を適用した学習装置のさらに他の例を示すブロック図である。

図28は、本発明に学習装置を構成する予測フィルタを示すブロック図である。

図29は、図27に示す学習装置の処理を説明するフローチャートである。

図30は、本発明を適用した伝送システムを示すブロック図である。

図31は、本発明が適用された携帯電話機を示すブロック図である。

図32は、携帯電話機を構成する受信部を示すブロック図である。

図33は、本発明を適用した学習装置の他を示すブロック図である。

図34は、教師データと生徒データを示す図である。

発明を実施するための最良の形態

以下、本発明の実施の形態を図面を参照して詳細に説明する。

本発明を適用した音声合成装置は、図3に示すような構成を備え、音声合成フィルタ44に与える残差信号と線形予測係数を、それぞれベクトル量子化等によってコード化した残差コードとAコードが多重化されたコードデータが供給されるようになっており、その残差コードとAコードから、それぞれ残差信号と線形予測係数を復号し、音声合成フィルタ44に与えることで、合成音が生成されるようになっている。この音声合成装置では、音声合成フィルタ44で生成された合成音と、学習により求めたタップ係数を用いた予測演算を行うことにより、その合成音の音質を向上させた高音質の音声を求めて出力する。

本発明を適用した図3の音声合成装置では、クラス分類適応処理を利用して、合成音が、真の高音質の音声（の予測値）に復号される。

クラス分類適応処理は、クラス分類処理と適応処理とからなり、クラス分類処理によって、データを、その性質に基づいてクラス分けし、各クラスごとに適応処理を施すものであり、適応処理は、以下の手法のものである。

即ち、適応処理では、例えば、合成音と、所定のタップ係数との線形結合により、真の高音質の音声の予測値が求められる。

具体的には、例えば、いま、真の高音質の音声（のサンプル値）を教師データとするとともに、その真の高音質の音声を、CELP方式によって、Lコード、Gコード、Iコード、及びAコードに符号化し、それらのコードを、前述した図2に示した受信部で復号することにより得られる合成音を生徒データとして、教

師データである高音質の音声 y の予測値 $E[y]$ を、幾つかの合成音（のサンプル値） x_1, x_2, \dots の集合と、所定のタップ係数 w_1, w_2, \dots の線形結合により規定される線形 1 次結合モデルにより求めることを考える。この場合、予測値 $E[y]$ は、次式で表すことができる。

$$E[y] = w_1 x_1 + w_2 x_2 + \dots \quad \dots \quad (6)$$

式 (6) を一般化するために、タップ係数 w_i の集合である行列 W 、生徒データ x_{ij} の集合である行列 X 、及び予測値 $E[y_i]$ の集合である行列 Y' を、

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1J} \\ x_{21} & x_{22} & \dots & x_{2J} \\ \dots & \dots & \dots & \dots \\ x_{I1} & x_{I2} & \dots & x_{IJ} \end{bmatrix},$$

$$W = \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_J \end{bmatrix}, Y' = \begin{bmatrix} E[y_1] \\ E[y_2] \\ \dots \\ E[y_I] \end{bmatrix}$$

で定義すると、次のような観測方程式が成立する。

$$XW = Y' \quad \dots \quad (7)$$

ここで、行列 X の成分 x_{ij} は、 i 件目の生徒データの集合（ i 件目の教師データ y_i の予測に用いる生徒データの集合）の中の j 番目の生徒データを意味し、行列 W の成分 w_j は、生徒データの集合の中の j 番目の生徒データとの積が演算されるタップ係数を表す。また、 y_i は、 i 件目の教師データを表し、従って、 $E[y_i]$ は、 i 件目の教師データの予測値を表す。なお、式 (6) の左辺における y は、行列 Y' の成分 y_i のサフィックス i を省略したものであり、式 (6) の右辺における x_1, x_2, \dots も、行列 X の成分 x_{ij} のサフィックス i を省略したものである。

この観測方程式に最小自乗法を適用して、真の高音質の音声 y に近い予測値 $E[y]$ を求めることを考える。この場合、教師データとなる真の高音質の音声 y の集合である行列 Y 、及び高音質の音声 y に対する予測値 $E[y]$ の残差 e の集

合でなる行列Eを、

$$E = \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_I \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_I \end{bmatrix}$$

で定義すると、式(7)から、次のような残差方程式が成立する。

$$XW = Y + E \quad \dots \quad (8)$$

この場合、真の高音質の音声yに近い予測値E[y]を求めるためのタップ係数w_jは、自乗誤差

$$\sum_{i=1}^I e_i^2$$

を最小にすることで求めることができる。

上述の自乗誤差をタップ係数w_jで微分したものが0になる場合、即ち、次式を満たすタップ係数w_jが、真の高音質の音声yに近い予測値E[y]を求めるため最適値ということになる。

$$e_1 \frac{\partial e_1}{\partial w_j} + e_2 \frac{\partial e_2}{\partial w_j} + \dots + e_I \frac{\partial e_I}{\partial w_j} = 0 \quad (j = 1, 2, \dots, J)$$

そこで、まず、式(8)を、タップ係数w_jで微分することにより、次式が成立する。

$$\frac{\partial e_1}{\partial w_1} = x_{11}, \frac{\partial e_1}{\partial w_2} = x_{12}, \dots, \frac{\partial e_I}{\partial w_J} = x_{IJ} \quad (i = 1, 2, \dots, I)$$

式(9)及び(10)より、式(11)が得られる。

$$\sum_{i=1}^I e_i x_{i1} = 0, \sum_{i=1}^I e_i x_{i2} = 0, \dots, \sum_{i=1}^I e_i x_{iJ} = 0 \quad \dots (11)$$

さらに、式(8)の残差方程式における生徒データ x_{ij} 、タップ係数 w_j 、教師データ y_i 、及び誤差 e_i の関係を考慮すると、式(11)から、次のような正規方程式を得ることができる。

$$\begin{cases} \left(\sum_{i=1}^I X_{i1} X_{i1} \right) W_1 + \left(\sum_{i=1}^I X_{i1} X_{i2} \right) W_2 + \dots + \left(\sum_{i=1}^I X_{i1} X_{iJ} \right) W_J = \left(\sum_{i=1}^I X_{i1} Y_i \right) \\ \left(\sum_{i=1}^I X_{i2} X_{i1} \right) W_1 + \left(\sum_{i=1}^I X_{i2} X_{i2} \right) W_2 + \dots + \left(\sum_{i=1}^I X_{i2} X_{iJ} \right) W_J = \left(\sum_{i=1}^I X_{i2} Y_i \right) \\ \dots \\ \left(\sum_{i=1}^I X_{iJ} X_{i1} \right) W_1 + \left(\sum_{i=1}^I X_{iJ} X_{i2} \right) W_2 + \dots + \left(\sum_{i=1}^I X_{iJ} X_{iJ} \right) W_J = \left(\sum_{i=1}^I X_{iJ} Y_i \right) \end{cases}$$

... (12)

なお、式(12)に示した正規方程式は、行列（共分散行列）A及びベクトルvを、

$$A = \begin{pmatrix} \sum_{i=1}^I X_{i1} X_{i1} & \sum_{i=1}^I X_{i1} X_{i2} & \dots & \sum_{i=1}^I X_{i1} X_{iJ} \\ \sum_{i=1}^I X_{i2} X_{i1} & \sum_{i=1}^I X_{i2} X_{i2} & \dots & \sum_{i=1}^I X_{i2} X_{iJ} \\ \dots & & & \\ \sum_{i=1}^I X_{iJ} X_{i1} & \sum_{i=1}^I X_{iJ} X_{i2} & \dots & \sum_{i=1}^I X_{iJ} X_{iJ} \end{pmatrix}$$

$$v = \begin{pmatrix} \sum_{i=1}^I X_{i1} Y_i \\ \sum_{i=1}^I X_{i2} Y_i \\ \vdots \\ \sum_{i=1}^I X_{iJ} Y_i \end{pmatrix}$$

で定義するとともに、ベクトルWを、数1で示したように定義すると、式

$$A W = v \quad \cdots (13)$$

で表すことができる。

式(12)における各正規方程式は、生徒データ x_i 及び教師データ y_i のセットを、ある程度の数だけ用意することで、求めるべきタップ係数 w_j の数 J と同じ数だけたてることができ、従って、式(13)を、ベクトルWについて解くことで（但し、式(13)を解くには、式(13)における行列Aが正則である必要がある）、最適なタップ係数（ここでは、自乗誤差を最小にするタップ係数） w_j を求めることができる。なお、式(13)を解くにあたっては、例えば、掃き出し法（Gauss-Jordanの消去法）などを用いることが可能である。

以上のようにして、最適なタップ係数 w_j を求めておき、さらに、そのタップ係数 w_j を用い、式(6)により、真の高音質の音声 y に近い予測値 $E[y]$ を求めるのが適応処理である。

なお、教師データとして、高いサンプリング周波数でサンプリングした音声信号、又は多ビットを割り当てた音声信号を用いるとともに、生徒データとして、その教師データとしての音声信号を間引いたり、低ビットで再量子化した音声信号をCELP方式により符号化し、その符号化結果を復号して得られる合成音を用いた場合、タップ係数としては、高いサンプリング周波数でサンプリングした音声信号、又は多ビットを割り当てた音声信号を生成するのに、予測誤差が、統計的に最小となる高音質の音声が得られることになる。この場合、より高音質の合成音を得ることが可能となる。

図3の音声合成装置では、以上のようなクラス分類適応処理により、Aコードと残差コードであるコードデータを、高音質の音声に復号するようになっている。

即ち、デマルチブレクサ（DEMUX）41には、コードデータが供給されるようになっており、デマルチブレクサ41は、そこに供給されるコードデータから、フレームごとのAコードと残差コードを分離する。そして、デマルチブレクサは、Aコードを、フィルタ係数復号器42及びタップ生成部46に供給し、残差コードを、残差コードブック記憶部43及びタップ生成部46に供給する。

ここで、図3におけるコードデータに含まれるAコードと残差コードは、音声

をLPC分析して得られる線形予測係数と残差信号を、所定のコードブックを用いて、それぞれベクトル量子化することにより得られるコードとなっている。

フィルタ係数復号器42は、デマルチプレクサ41から供給されるフレームごとのAコードを、そのAコードを得るときに用いられたのと同一のコードブックに基づいて、線形予測係数に復号し、音声合成フィルタ44に供給する。

残差コードブック記憶部43は、デマルチプレクサ41から供給されるフレームごとの残差コードを、その残差コードを得るときに用いられたのと同一のコードブックに基づいて、残差信号に復号し、音声合成フィルタ44に供給する。

音声合成フィルタ44は、例えば、図1の音声合成フィルタ29と同様に、IIR型のディジタルフィルタで、フィルタ係数復号器42からの線形予測係数をIIRフィルタのタップ係数とするとともに、残差コードブック記憶部43からの残差信号を入力信号として、その入力信号のフィルタリングを行うことにより、合成音を生成し、タップ生成部45に供給する。

タップ生成部45は、音声合成フィルタ44から供給される合成音のサンプル値から、後述する予測部49における予測演算に用いられる予測タップとなるものを抽出する。即ち、タップ生成部45は、例えば、高音質の音声の予測値を求めようとしているフレームである注目フレームの合成音のサンプル値すべてを、予測タップとする。そして、タップ生成部45は、予測タップを、予測部49に供給する。

タップ生成部46は、デマルチプレクサ41から供給されるフレーム又はサブフレームごとのAコード及び残差コードから、クラスタップとなるものを抽出する。即ち、タップ生成部46は、例えば、注目フレームのAコード及び残差コードすべてを、クラスタップとする。タップ生成部46は、クラスタップを、クラス分類部47に供給する。

ここで、予測タップやクラスタップの構成パターンは、上述したパターンのものに限定されるものではない。

なお、タップ生成部46では、Aコードや残差コードの他、フィルタ係数復号器42が出力する線形予測係数や、残差コードブック記憶部43が出力する残差信号、さらには、音声合成フィルタ44が出力する合成音等の中からも、クラス

タップを抽出するようにすることができる。

クラス分類部47は、タップ生成部46からのクラスタップに基づき、注目している注目フレームの音声（のサンプル値）をクラス分類し、その結果得られるクラスに対応するクラスコードを、係数メモリ48に出力する。

ここで、クラス分類部47には、例えば、クラスタップとしての注目フレームのAコード及び残差コードを構成するビットの系列そのものを、クラスコードとして出力させることができある。

係数メモリ48は、後述する図6の学習装置において学習処理が行われることにより得られる、クラスごとのタップ係数を記憶しており、クラス分類部47が出力するクラスコードに対応するアドレスに記憶されているタップ係数を、予測部49に出力する。

ここで、各フレームについて、Nサンプルの高音質の音声が求められるとすると、注目フレームについて、Nサンプルの音声を、式(6)の予測演算によって求めるには、Nセットのタップ係数が必要である。従って、この場合は、係数メモリ48には、1つのクラスコードに対応するアドレスに対して、Nセットのタップ係数が記憶されている。

予測部49は、タップ生成部45が出力する予測タップと、係数メモリ48が出力するタップ係数とを取得し、その予測タップとタップ係数とを用いて、式(6)に示した線形予測演算（積和演算）を行い、注目フレームの高音質の音声の予測値を求めて、D/A変換部50に出力する。

ここで、係数メモリ48は、上述したように、注目フレームの音声のNサンプルそれぞれを求めるためのNセットのタップ係数を出力するが、予測部49は、各サンプル値を、予測タップと、そのサンプル値に対応するタップ係数のセットとを用いて、式(6)の積和演算を行う。

D/A変換部50は、予測部49からの音声（の予測値）を、デジタル信号からアナログ信号にD/A変換し、スピーカ51に供給して出力させる。

次に、図4は、図3の音声合成フィルタ44の構成例を示している。

図4において、音声合成フィルタ44は、P次の線形予測係数を用いるものとなっており、従って、1つの加算器61、P個の遅延回路(D)62乃至62P、

及びP個の乗算器63₁乃至63_Pから構成されている。

乗算器63₁乃至63_Pには、それぞれ、フィルタ係数復号器42から供給されるP次の線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_P$ がセットされ、これにより、音声合成フィルタ44では、式(4)にしたがって演算が行われ、合成音が生成される。

即ち、残差コードブック記憶部43が出力する残差信号eは、加算器61を介して、遅延回路62₁に供給され、遅延回路62₁は、そこへの入力信号を、残差信号の1サンプル分だけ遅延して、後段の遅延回路62_{p+1}に出力するとともに、演算器63_pに出力する。乗算器63_pは、遅延回路62_pの出力と、そこにセットされた線形予測係数 α_p とを乗算し、その乗算値を、加算器61に出力する。

加算器61は、乗算器63₁乃至63_Pの出力すべてと、残差信号eとを加算し、その加算結果を、遅延回路62₁に供給する他、音声合成結果(合成音)として出力する。

次に、図5のフローチャートを参照して、図3の音声合成装置の音声合成処理について説明する。

デマルチブレクサ41は、そこに供給されるコードデータから、フレームごとのAコードと残差コードを順次分離し、それを、フィルタ係数復号器42と残差コードブック記憶部43に供給する。さらに、デマルチブレクサ41は、Aコード及び残差コードを、タップ生成部46に供給する。

フィルタ係数復号器42は、デマルチブレクサ41から供給されるフレームごとのAコードを、線形予測係数に順次復号し、音声合成フィルタ44に供給する。また、残差コードブック記憶部43は、デマルチブレクサ41から供給されるフレームごとの残差コードを、残差信号に順次復号し、音声合成フィルタ44に供給する。

音声合成フィルタ44では、そこに供給される残差信号及び線形予測係数を用いて、前述した式(4)の演算が行われることにより、注目フレームの合成音が生成される。この合成音は、タップ生成部45に供給される。

タップ生成部45は、そこに供給される合成音のフレームを、順次、注目フレームとし、ステップS1において、音声合成フィルタ44から供給される合成音のサンプル値から、予測タップを生成し、予測部49に出力する。さらに、ステ

ップS1では、タップ生成部46が、デマルチフレクサ41から供給されるAコード及び残差コードから、クラスタップを生成し、クラス分類部47に出力する。

ステップS2に進み、クラス分類部47は、タップ生成部46から供給されるクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、係数メモリ48に供給して、ステップS3に進む。

ステップS3では、係数メモリ48は、クラス分類部47から供給されるクラスコードに対応するアドレスから、タップ係数を読み出し、予測部49に供給する。

ステップS4に進み、予測部49は、係数メモリ48が出力するタップ係数を取得し、そのタップ係数と、タップ生成部45からの予測タップとを用いて、式(6)に示した積和演算を行い、注目フレームの高音質の音声の予測値を得る。この高音質の音声は、予測部49からD/A変換部50を介して、スピーカ51に供給されて出力される。

予測部49において、注目フレームの高音質の音声が得られた後は、ステップS5に進み、まだ、注目フレームとして処理すべきフレームがあるかどうかが判定される。ステップS5において、まだ、注目フレームとして処理すべきフレームがあると判定された場合、ステップS1に戻り、次に注目フレームとすべきフレームを、新たに注目フレームとして、以下、同様の処理を繰り返す。また、ステップS5において、注目フレームとして処理すべきフレームがないと判定された場合、音声合成処理を終了する。

次に、図3の係数メモリ48に記憶させるタップ係数の学習処理を行う学習装置の一例を図6を参照して説明する。

図6に示す学習装置には、学習用のディジタル音声信号が、所定のフレーム単位で供給されるようになっており、この学習用のディジタル音声信号は、LPC分析部71及び予測フィルタ74に供給される。さらに、学習用のディジタル音声信号は、教師データとして、正規方程式加算回路81にも供給される。

LPC分析部71は、そこに供給される音声信号のフレームを、順次、注目フレームとし、その注目フレームの音声信号をLPC分析することで、P次の線形予測係数を求め、予測フィルタ74及びベクトル量子化部72に供給する。

ベクトル量子化部72は、線形予測係数を要素とするコードベクトルとコードとを対応付けたコードブックを記憶しており、そのコードブックに基づいて、LPC分析部71からの注目フレームの線形予測係数で構成される特徴ベクトルをベクトル量子化し、そのベクトル量子化の結果得られるAコードを、フィルタ係数復号器73及びタップ生成部79に供給する。

フィルタ係数復号器73は、ベクトル量子化部72が記憶しているのと同一のコードブックを記憶しており、そのコードブックに基づいて、ベクトル量子化部72からのAコードを、線形予測係数に復号し、音声合成フィルタ77に供給する。ここで、図3のフィルタ係数復号器42は、図6のフィルタ係数復号器73と同様に構成されている。

予測フィルタ74は、そこに供給される注目フレームの音声信号と、LPC分析部71からの線形予測係数を用いて、例えば、前述した式(1)にしたがった演算を行うことにより、注目フレームの残差信号を求め、ベクトル量子化部75に供給する。

即ち、式(1)における s_n と e_n のZ変換を、SとEとそれぞれ表すと、式(1)は、次式のように表すことができる。

$$E = (1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots + \alpha_p z^{-p}) S \quad \dots \quad (14)$$

式(14)から、残差信号eを求める予測フィルタ74は、FIR(Finite Impulse Response)型のデジタルフィルタで構成することができる。

即ち、図7は、予測フィルタ74の構成例を示している。

予測フィルタ74には、LPC分析部71から、P次の線形予測係数が供給されるようになっており、従って、予測フィルタ74は、P個の遅延回路(D)91, 乃至91p, P個の乗算器921, 乃至92p, 及び1つの加算器93から構成されている。

乗算器921, 乃至92pには、それぞれ、LPC分析部71から供給されるP次の線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ がセットされる。

一方、注目フレームの音声信号sは、遅延回路911と加算器93に供給される。遅延回路91pは、そこへの入力信号を、残差信号の1サンプル分だけ遅延して、後段の遅延回路91p+1に出力するとともに、演算器92pに出力する。乗算

器92は、遅延回路91の出力と、そこにセットされた線形予測係数 α_2 とを乗算し、その乗算値を、加算器93に出力する。

加算器93は、乗算器92乃至92₂の出力すべてと、音声信号sとを加算し、その加算結果を、残差信号eとして出力する。

図6に戻り、ベクトル量子化部75は、残差信号のサンプル値を要素とするコードベクトルとコードとを対応付けたコードブックを記憶しており、そのコードブックに基づいて、予測フィルタ74からの注目フレームの残差信号のサンプル値で構成される残差ベクトルをベクトル量子化し、そのベクトル量子化の結果得られる残差コードを、残差コードブック記憶部76及びタップ生成部79に供給する。

残差コードブック記憶部76は、ベクトル量子化部75が記憶しているのと同一のコードブックを記憶しており、そのコードブックに基づいて、ベクトル量子化部75からの残差コードを、残差信号に復号し、音声合成フィルタ77に供給する。ここで、図3の残差コードブック記憶部43は、図6の残差コードブック記憶部76と同様に構成されている。

音声合成フィルタ77は、図3の音声合成フィルタ44と同様に構成されるIIRフィルタで、フィルタ係数復号器73からの線形予測係数をIIRフィルタのタップ係数とともに、残差コードブック記憶部75からの残差信号を入力信号として、その入力信号のフィルタリングを行うことにより、合成音を生成し、タップ生成部78に供給する。

タップ生成部78は、図3のタップ生成部45における場合と同様に、音声合成フィルタ77から供給される線形予測係数から、予測タップを構成し、正規方程式加算回路81に供給する。タップ生成部79は、図3のタップ生成部46における場合と同様に、ベクトル量子化部72と75からそれぞれ供給されるAコードと残差コードから、クラスタップを構成し、クラス分類部80に供給する。

クラス分類部80は、図3のクラス分類部47における場合と同様に、そこに供給されるクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路81に供給する。

正規方程式加算回路81は、教師データとしての注目フレームの高音質の音声

である学習用の音声と、タップ生成部 78 からの生徒データとしての予測タップを構成する音声合成フィルタ 77 の合成音出力を対象とした足し込みを行う。

即ち、正規方程式加算回路 81 は、クラス分類部 80 から供給されるクラスコードに対応するクラスごとに、予測タップ（生徒データ）を用い、式（13）の行列 A における各コンポーネントとなっている、生徒データどうしの乗算 ($x_{in} \cdot x_{im}$) と、サメーション (Σ) に相当する演算を行う。

さらに、正規方程式加算回路 81 は、やはり、クラス分類部 80 から供給されるクラスコードに対応するクラスごとに、生徒データ、即ち、予測タップを構成する、音声合成フィルタ 77 から出力される合成音のサンプル値及び教師データ、即ち、注目フレームの高音質の音声のサンプル値を用い、式（13）のベクトル v における各コンポーネントとなっている、生徒データと教師データの乗算 ($x_{in} \cdot y_i$) と、サメーション (Σ) に相当する演算を行う。

正規方程式加算回路 81 は、以上の足し込みを、そこに供給される学習用の音声のフレームすべてを注目フレームとして行い、これにより、各クラスについて、式（13）に示した正規方程式をたてる。

タップ係数決定回路 82 は、正規方程式加算回路 81 においてクラスごとに生成された正規方程式を解くことにより、クラスごとに、タップ係数を求め、係数メモリ 83 の、各クラスに対応するアドレスに供給する。

なお、学習用の音声信号として用意した音声信号によっては、正規方程式加算回路 81 において、タップ係数を求めるのに必要な数の正規方程式が得られないクラスが生じる場合があり得るが、タップ係数決定回路 82 は、そのようなクラスについては、例えば、デフォルトのタップ係数を出力する。

係数メモリ 83 は、タップ係数決定回路 82 から供給されるクラスごとのタップ係数を、そのクラスに対応するアドレスに記憶する。

次に、図 8 のフローチャートを参照して、図 6 の学習装置の学習処理について説明する。

学習装置には、学習用の音声信号が供給され、この学習用の音声信号は、LPC 分析部 71 及び予測フィルタ 74 に供給されるとともに、教師データとして、正規方程式加算回路 81 に供給される。そして、ステップ S11 において、学習

用の音声信号から、生徒データが生成される。

即ち、LPC分析部71は、学習用の音声信号のフレームを、順次、注目フレームとし、その注目フレームの音声信号をLPC分析することで、P次の線形予測係数を求め、ベクトル量子化部72に供給する。ベクトル量子化部72は、LPC分析部71からの注目フレームの線形予測係数で構成される特徴ベクトルをベクトル量子化し、そのベクトル量子化の結果得られるAコードを、フィルタ係数復号器73及びタップ係数生成部79に供給する。フィルタ係数復号器73は、ベクトル量子化部72からのAコードを、線形予測係数に復号し、その線形予測係数を、音声合成フィルタ77に供給する。

一方、LPC分析部71から注目フレームの線形予測係数を受信した予測フィルタ74は、その線形予測係数と、注目フレームの学習用の音声信号とを用いて、式(1)にしたがった演算を行うことにより、注目フレームの残差信号を求め、ベクトル量子化部75に供給する。ベクトル量子化部75は、予測フィルタ74からの注目フレームの残差信号のサンプル値で構成される残差ベクトルをベクトル量子化し、そのベクトル量子化の結果得られる残差コードを、残差コードブック記憶部76及びタップ生成部79に供給する。残差コードブック記憶部76は、ベクトル量子化部75からの残差コードを、残差信号に復号し、音声合成フィルタ77に供給する。

以上のようにして、音声合成フィルタ77は、線形予測係数と残差信号を受信すると、その線形予測係数と残差信号を用いて音声合成を行い、その結果得られる合成音を、生徒データとして、タップ生成部78に出力する。

そして、ステップS12に進み、タップ生成部78が、音声合成フィルタ77から供給される合成音から、予測タップを生成するとともに、タップ生成部79が、ベクトル量子化部72からのAコードと、ベクトル量子化部75からの残差コードから、クラスタップを生成する。予測タップは、正規方程式加算回路81に供給され、クラスタップは、クラス分類部80に供給される。

その後、ステップS13において、クラス分類部80が、タップ生成部79からのクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路81に供給する。

ステップS14に進み、正規方程式加算回路81は、クラス分類部80から供給されるクラスについて、そこに供給される教師データとしての注目フレームの高音質の音声のサンプル値、及びタップ生成部78からの生徒データとしての予測タップ（を構成する合成音のサンプル値）を対象とした、式（13）の行列Aとベクトルvの、上述したような足し込みを行い、ステップS15に進む。

ステップS15では、まだ、注目フレームとして処理すべきフレームの学習用の音声信号があるかどうかが判定される。ステップS15において、まだ、注目フレームとして処理すべきフレームの学習用の音声信号があると判定された場合、ステップS11に戻り、次のフレームを新たに注目フレームとして、以下、同様の処理が繰り返される。

ステップS15において、注目フレームとして処理すべきフレームの学習用の音声信号がないと判定された場合、即ち、正規方程式加算回路81において、各クラスについて、正規方程式が得られた場合、ステップS16に進み、タップ係数決定回路82は、各クラスごとに生成された正規方程式を解くことにより、各クラスごとに、タップ係数を求め、係数メモリ83の、各クラスに対応するアドレスに供給して記憶させ、処理を終了する。

以上のようにして、係数メモリ83に記憶された各クラスごとのタップ係数が、図3の係数メモリ48に記憶されている。

従って、図3の係数メモリ48に記憶されたタップ係数は、線形予測演算を行うことにより得られる高音質の音声の予測値の予測誤差、ここでは、自乗誤差が、統計的に最小になるように学習を行うことにより求められたものであるから、図3の予測部49が出力する音声は、音声合成フィルタ44で生成された合成音の歪みが低減（解消）された、高音質のものとなる。

なお、図3の音声合成装置において、上述したように、例えば、タップ生成部46に、線形予測係数や残差信号等の中からも、クラスタップを抽出させるようとする場合には、図6のタップ生成部79にも、フィルタ係数復号器73が出力する線形予測係数や、残差コードブック記憶部76が出力する残差信号の中から、同様のクラスタップを抽出せしめる必要がある。但し、線形予測係数等からも、クラスタップを抽出する場合には、タップ数が多くなることから、クラ

ス分類は、例えば、クラスタップをベクトル量子化等によって圧縮することにより行うのが望ましい。なお、残差コード及びAコードだけからクラス分類を行う場合には、残差コードとAコードのビット列の並びを、そのままクラスコードとすることができることから、クラス分類処理に要する負担を軽減することができる。

次に、本発明を適用した伝送システムの一例を、図9を参照して説明する。ここで、システムとは、複数の装置が論理的に集合した物をいい、各構成の装置が同一筐体中にあるか否かは問わない。

図9に示す伝送システムでは、携帯電話機101₁と101₂が、基地局102₁と102₂それぞれとの間で、無線による送受信を行うとともに、基地局102₁と102₂、それぞれが、交換局103との間で送受信を行うことにより、最終的には、携帯電話機101₁と101₂との間において、基地局102₁及び102₂、並びに交換局103を介して、音声の送受信を行うことができるようになっている。なお、基地局102₁と102₂は、同一の基地局であってもよいし、異なる基地局であってもよい。

ここで、以下、特に区別する必要がない限り、携帯電話機101₁と101₂を、携帯電話機101と記述する。

図10は、図9に示す携帯電話機101の構成例を示している。

アンテナ111は、基地局102₁又は102₂からの電波を受信し、その受信信号を、変復調部112に供給するとともに、変復調部112からの信号を、電波で、基地局102₁又は102₂に送信する。変復調部112は、アンテナ111からの信号を復調し、その結果得られる、図1で説明したようなコードデータを、受信部114に供給する。また、変復調部112は、送信部113から供給される、図1で説明したようなコードデータを変調し、その結果得られる変調信号を、アンテナ111に供給する。送信部113は、図1に示した送信部と同様に構成され、そこに入力されるユーザの音声を、コードデータに符号化して、変復調部112に供給する。受信部114は、変復調部112からのコードデータを受信し、そのコードデータから、図3の音声合成装置における場合と同様の高音質の音声を復号して出力する。

即ち、図11は、図10の受信部114の構成例を示している。なお、図中、図2における場合と対応する部分については、同一の符号を付してあり、以下では、その説明は、適宜省略する。

タップ生成部121には、音声合成フィルタ29が出力する合成音が供給されるようになっており、タップ生成部121は、その合成音から、予測タップとするもの（サンプル値）を抽出し、予測部125に供給する。

タップ生成部122には、チャネルデコーダ21が出力する、フレーム又はサブフレームごとのLコード、Gコード、Iコード、及びAコードが供給されるようになっている。さらに、タップ生成部122には、演算器28から残差信号が供給されるとともに、フィルタ係数復号器25から線形予測係数が供給されるようになっている。タップ生成部122は、そこに供給されるLコード、Gコード、Iコード、及びAコード、さらには、残差信号及び線形予測係数から、クラスタップとするものを抽出し、クラス分類部123に供給する。

クラス分類部123は、タップ生成部122から供給されるクラスタップに基づいて、クラス分類を行い、そのクラス分類結果としてのクラスコードを、係数メモリ124に供給する。

ここで、Lコード、Gコード、Iコード、及びAコード、並びに残差信号及び線形予測係数から、クラスタップを構成し、このクラスタップに基づいてクラス分類を行うと、そのクラス分類の結果得られるクラス数が膨大な数になることがある。そこで、クラス分類部123では、例えば、Lコード、Gコード、Iコード、及びAコード、並びに残差信号及び線形予測係数を要素とするベクトルをベクトル量子化して得られるコードを、クラス分類結果として出力するようにすることができる。

係数メモリ124は、後述する図12の学習装置において学習処理が行われることにより得られる、クラスごとのタップ係数を記憶しており、クラス分類部123が出力するクラスコードに対応するアドレスに記憶されているタップ係数を、予測部125に供給する。

予測部125は、図3の予測部49と同様に、タップ生成部121が出力する予測タップと、係数メモリ124が出力するタップ係数とを取得し、その予測タ

ップとタップ係数とを用いて、式(6)に示した線形予測演算を行う。これにより、予測部125は、注目フレームの高音質の音声(の予測値)を求めて、D/A変換部30に供給する。

以上のように構成される受信部114では、基本的には、図5に示したフローチャートにしたがった処理と同様の処理が行われることで、高音質の合成音が、音声の復号結果として出力される。

即ち、チャネルデコーダ21は、そこに供給されるコードデータから、Lコード、Gコード、Iコード、Aコードを分離し、それぞれを、適応コードブック記憶部22、ゲイン復号器23、励起コードブック記憶部24、フィルタ係数復号器25に供給する。さらに、Lコード、Gコード、Iコード、及びAコードは、タップ生成部122にも供給される。

適応コードブック記憶部22、ゲイン復号器23、励起コードブック記憶部24、演算器26乃至28では、図1の適応コードブック記憶部9、ゲイン復号器10、励起コードブック記憶部11、演算器12乃至14における場合と同様の処理が行われ、これにより、Lコード、Gコード、及びIコードが、残差信号eに復号される。この残差信号は、音声合成フィルタ29及びタップ生成部122に供給される。

フィルタ係数復号器25は、図1で説明したように、そこに供給されるAコードを、線形予測係数に復号し、音声合成フィルタ29及びタップ生成部122に供給する。音声合成フィルタ29は、演算器28からの残差信号と、フィルタ係数復号器25からの線形予測係数を用いて音声合成を行い、その結果得られる合成音を、タップ生成部121に供給する。

タップ生成部121は、音声合成フィルタ29が output する合成音のフレームを注目フレームとし、ステップS1において、その注目フレームの合成音から、予測タップを生成し、予測部125に供給する。さらに、ステップS1では、タップ生成部122は、そこに供給されるLコード、Gコード、Iコード、及びAコード、並びに残差信号及び線形予測係数から、クラスタップを生成し、クラス分類部123に供給する。

ステップS2に進み、クラス分類部123は、タップ生成部122から供給さ

れるクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、係数メモリ124に供給して、ステップS3に進む。

ステップS3では、係数メモリ124は、クラス分類部123から供給されるクラスコードに対応するアドレスから、タップ係数を読み出し、予測部125に供給する。

ステップS4に進み、予測部125は、係数メモリ124が出力する残差信号についてのタップ係数を取得し、そのタップ係数と、タップ生成部121からの予測タップとを用いて、式(6)に示した積和演算を行い、注目フレームの高音質の音声の予測値を得る。

以上のようにして得られた高音質の音声は、予測部125から、D/A変換部30を介して、スピーカ31に供給され、これにより、スピーカ31からは、高音質の音声が 출력される。

ステップS4の処理後は、ステップS5に進み、まだ、注目フレームとして処理すべきフレームがあるかどうかが判定され、あると判定された場合、ステップS1に戻り、次に注目フレームとすべきフレームを、新たに注目フレームとして、以下、同様の処理を繰り返す。また、ステップS5において、注目フレームとして処理すべきフレームがないと判定された場合、処理を終了する。

次に、図11の係数メモリ124に記憶させるタップ係数の学習処理を行う学習装置の一例を図12に示す。

図12に示す学習装置において、マイクロフォン201乃至コード決定部215は、図1のマイクロフォン1乃至コード決定部15とそれぞれ同様に構成される。マイクロフォン1には、学習用の音声信号が入力されるようになっており、従って、マイクロフォン201乃至コード決定部215では、その学習用の音声信号に対して、図1における場合と同様の処理が施される。

タップ生成部131には、自乗誤差最小判定部208において自乗誤差が最小になったと判定されたときの音声合成フィルタ206が出力する合成音が供給される。また、タップ生成部132には、コード決定部215が、自乗誤差最小判定部208から確定信号を受信したときに出力するLコード、Gコード、Iコード、及びAコードが供給される。さらに、タップ生成部132には、ベクトル量

子化部 205 が output する、LPC 分析部 204 で得られた線形予測係数のベクトル量子化結果としての A コードに対応するコードベクトル（セントロイドベクトル）の要素となっている線形予測係数と、自乗誤差最小判定部 208 において自乗誤差が最小になったと判定されたときの演算器 214 が output する残差信号も供給される。また、正規方程式加算回路 134 には、A/D 変換部 202 が output する音声が、教師データとして供給される。

タップ生成部 131 は、音声合成フィルタ 206 が output する合成音から、図 11 のタップ生成部 121 と同一の予測タップを構成し、生徒データとして、正規方程式加算回路 134 に供給する。

タップ生成部 132 は、コード決定部 215 から供給される L コード、G コード、I コード、及び A コード、並びに、ベクトル量子化部 205 から供給される線形予測係数、及び演算器 214 から供給される残差信号から、図 11 のタップ生成部 122 と同一のクラスタップを構成し、クラス分類部 133 に供給する。

クラス分類部 133 は、タップ生成部 132 からのクラスタップに基づいて、図 11 のクラス分類部 123 における場合と同様のクラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路 134 に供給する。

正規方程式加算回路 134 は、A/D 変換部 202 からの音声を、教師データとして受信するとともに、タップ生成部 131 からの予測タップを、生徒データとして受信し、その教師データ及び生徒データを対象として、クラス分類部 133 からのクラスコードごとに、図 6 の正規方程式加算回路 81 における場合と同様の足し込みを行うことにより、各クラスについて、式 (13) に示した正規方程式をたてる。

タップ係数決定回路 135 は、正規方程式加算回路 134 においてクラスごとに生成された正規方程式を解くことにより、クラスごとに、タップ係数を求め、係数メモリ 136 の、各クラスに対応するアドレスに供給する。

なお、学習用の音声信号として用意する音声信号によっては、正規方程式加算回路 134 において、タップ係数を求めるのに必要な数の正規方程式が得られないクラスが生じる場合があり得るが、タップ係数決定回路 135 は、そのようなクラスについては、例えば、デフォルトのタップ係数を output する。

係数メモリ 136 は、タップ係数決定回路 135 から供給されるクラスごとの線形予測係数と残差信号についてのタップ係数を記憶する。

以上のように構成される学習装置では、基本的には、図 8 に示したフローチャートにしたがった処理と同様の処理が行われることで、高音質の合成音を得るためにタップ係数が求められる。

学習装置には、学習用の音声信号が供給され、ステップ S11 では、その学習用の音声信号から、教師データと生徒データが生成される。

即ち、学習用の音声信号は、マイクロフォン 201 に入力され、マイクロフォン 201 乃至コード決定部 215 は、図 1 のマイクロフォン 1 乃至コード決定部 15 における場合とそれ同様の処理を行う。

その結果、A/D 変換部 202 で得られるデジタル信号の音声は、教師データとして、正規方程式加算回路 134 に供給される。また、自乗誤差最小判定部 208 において自乗誤差が最小になったと判定されたときに、音声合成フィルタ 206 が output する合成音は、生徒データとして、タップ生成部 131 に供給される。

さらに、ベクトル量子化部 205 が output する線形予測係数、自乗誤差最小判定部 208 において自乗誤差が最小になったと判定されたときに、コード決定部 215 が output する L コード、G コード、I コード、及び A コード、並びに演算器 214 が output する残差信号は、タップ生成部 132 に供給される。

その後、ステップ S12 に進み、タップ生成部 131 は、音声合成フィルタ 206 から生徒データとして供給される合成音のフレームを注目フレームとして、その注目フレームの合成音から、予測タップを生成し、正規方程式加算回路 134 に供給する。さらに、ステップ S12 では、タップ生成部 132 が、そこに供給される L コード、G コード、I コード、A コード、線形予測係数、及び残差信号から、クラスタップを生成し、クラス分類部 133 に供給する。

ステップ S12 の処理後は、ステップ S13 に進み、クラス分類部 133 が、タップ生成部 132 からのクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路 134 に供給する。

ステップ S14 に進み、正規方程式加算回路 134 は、A/D 変換器 202 か

らの教師データとしての注目フレームの高音質の音声である学習用の音声、及びタップ生成部 132 からの生徒データとしての予測タップを対象として、式(13)の行列 A とベクトル v の、上述したような足し込みを、クラス分類部 133 からのクラスコードごとに行い、ステップ S15 に進む。

ステップ S15 では、まだ、注目フレームとして処理すべきフレームがあるかどうかが判定される。ステップ S15において、まだ、注目フレームとして処理すべきフレームがあると判定された場合、ステップ S11 に戻り、次のフレームを新たに注目フレームとして、以下、同様の処理が繰り返される。

ステップ S15 において、注目フレームとして処理すべきフレームがないと判定された場合、即ち、正規方程式加算回路 134 において、各クラスについて、正規方程式が得られた場合、ステップ S16 に進み、タップ係数決定回路 135 は、各クラスごとに生成された正規方程式を解くことにより、各クラスごとに、タップ係数を求め、係数メモリ 136 の、各クラスに対応するアドレスに供給して記憶させ、処理を終了する。

以上のようにして、係数メモリ 136 に記憶された各クラスごとのタップ係数が、図 11 の係数メモリ 124 に記憶されている。

従って、図 11 の係数メモリ 124 に記憶されたタップ係数は、線形予測演算を行うことにより得られる高音質の音声予測値の予測誤差（自乗誤差）が、統計的に最小になるように学習を行うことにより求められたものであるから、図 11 の予測部 125 が出力する音声は、高音質のものとなる。

次に、上述した一連の処理は、ハードウェアにより行うこともできるし、ソフトウェアにより行うこともできる。一連の処理をソフトウェアによって行う場合には、そのソフトウェアを構成するプログラムが、汎用のコンピュータ等にインストールされる。

そこで、図 13 は、上述した一連の処理を実行するプログラムがインストールされるコンピュータの一実施の形態の構成例を示している。

プログラムは、コンピュータに内蔵されている記録媒体としてのハードディスク 305 や ROM 303 に予め記録しておくことができる。

あるいはまた、プログラムは、フロッピーディスク、CD-ROM(Compact D

isc Read Only Memory)、M O(Magneto Optical)ディスク、D V D(Digital Versatile Disc)、磁気ディスク、半導体メモリなどのリムーバブル記録媒体311に、一時的あるいは永続的に格納しておくことができる。このようなりムーバブル記録媒体311は、いわゆるパッケージソフトウェアとして提供することができる。

なお、プログラムは、上述したようなりムーバブル記録媒体311からコンピュータにインストールする他、ダウンロードサイトから、ディジタル衛星放送用の人工衛星を介して、コンピュータに無線で転送したり、L A N(Local Area Network)、インターネットといったネットワークを介して、コンピュータに有線で転送し、コンピュータでは、そのようにして転送されてくるプログラムを、通信部308で受信し、内蔵するハードディスク305にインストールすることができる。

コンピュータは、C P U(Central Processing Unit)302を内蔵している。C P U302には、バス301を介して、入出力インターフェース310が接続されており、C P U302は、入出力インターフェース310を介して、ユーザによって、キーボードや、マウス、マイクロフォン等で構成される入力部307が操作等されることにより指令が入力されると、それにしたがって、R O M(Read Only Memory)303に格納されているプログラムを実行する。あるいは、また、C P U302は、ハードディスク305に格納されているプログラム、衛星若しくはネットワークから転送され、通信部308で受信されてハードディスク305にインストールされたプログラム、又はドライブ309に装着されたリムーバブル記録媒体311から読み出されてハードディスク305にインストールされたプログラムを、R A M(Random Access Memory)304にロードして実行する。これにより、C P U302は、上述したフローチャートにしたがった処理、あるいは上述したブロック図の構成により行われる処理を行う。そして、C P U302は、その処理結果を、必要応じて、例えば、入出力インターフェース310を介して、L C D(Liquid Crystal Display)やスピーカ等で構成される出力部306から出力、あるいは、通信部308から送信、さらには、ハードディスク305に記録等させる。

ここで、コンピュータに各種の処理を行わせるためのプログラムを記述する処理ステップは、必ずしもフローチャートとして記載された順序に沿って時系列に処理する必要はなく、並列的あるいは個別に実行される処理、例えば、並列処理あるいはオブジェクトによる処理も含むものである。

また、プログラムは、1のコンピュータにより処理されるものであってもよいし、複数のコンピュータによって分散処理されるものであっても良い。さらに、プログラムは、遠方のコンピュータに転送されて実行されるものであってもよい。

なお、本発明において、学習用の音声信号として、どのようなものを用いるかについては、特に言及しなかったが、学習用の音声信号としては、人が発話した音声の他、例えば、曲（音楽）等を採用することが可能である。そして、上述したような学習処理によれば、学習用の音声信号として、人の発話を用いた場合には、そのような人の発話の音声の音質を向上させるようなタップ係数が得られ、曲を用いた場合には、曲の音質を向上させるようなタップ係数が得されることになる。

また、図11に示す例では、係数メモリ124には、タップ係数をあらかじめ記憶させておくようにしたが、係数メモリ124に記憶させるタップ係数は、携帯電話機101において、図9の基地局102、あるいは交換局103や、図示しないWWW(World Wide Web)サーバ等からダウンロードするようにすることができる。即ち、上述したように、タップ係数は、人の発話用や曲用等のように、ある種類の音声信号に適したものを、学習によって得ることができる。学習に用いる教師データ及び生徒データによっては、合成音の音質に差が生じるタップ係数を得ることができる。従って、そのような各種のタップ係数を、基地局102等に記憶させておき、ユーザには、自身の所望するタップ係数をダウンロードするようになることができる。そして、このようなタップ係数のダウンロードサービスは、無料で行うこともできるし、有料で行うこともできる。さらに、タップ係数のダウンロードサービスを有料で行う場合には、タップ係数のダウンロードに対する対価としての代金は、例えば、携帯電話機101の通話料等とともに請求するようになることができる。

係数メモリ124は、携帯電話機101に対して着脱可能なメモリカード等で

構成することができる。この場合、上述したような各種のタップ係数それぞれを記憶させた、異なるメモリカードを提供するようにすれば、ユーザは、場合に応じて、所望のタップ係数が記憶されたメモリカードを、携帯電話機 101 に装着して使用することが可能となる。

本発明は、例えば、V S E L P (Vector Sum Excited Liner Prediction)、P S I - C E L P (Pitch Synchronous Innovation CELP)、C S - A C E L P (Conjugate Structure Algebraic CELP) 等の C E L P 方式による符号化の結果得られるコードから合成音を生成する場合に、広く適用可能である。

また、本発明は、C E L P 方式による符号化の結果得られるコードから合成音を生成する場合に限らず、あるコードから、残差信号と線形予測係数を得て、合成音を生成する場合に、広く適用可能である。

上述した説明では、タップ係数を用いた線形 1 次予測演算によって、残差信号や線形予測係数の予測値を求めるようにしたが、この予測値は、その他、2 次以上の高次の予測演算によって求めることも可能である。

また、例えば、図 11 に示す受信部及び図 12 に示す学習装置では、クラスタップを、L コード、G コード、I コード、及び A コード の他、A コードから得られた線形予測係数や、L コード、G コード、及び I コードから得られた残差信号に基づいて生成するようにしたが、クラスタップは、その他、例えば、L コード、G コード、I コード、及び A コードだけから生成することも可能である。クラスタップは、4 種類の L コード、G コード、I コード、及び A コードのうちの 1 つだけ（又は複数）、即ち、例えば、I コードだけから生成することも可能である。例えば、クラスタップを、I コードだけから構成する場合においては、I コードそのものを、クラスコードとすることができる。ここで、V S E L P 方式では、I コードには、9 ビットが割り当てられており、従って、I コードを、そのままクラスコードとする場合、クラス数は、512 (= 2⁹) となる。なお、V S E L P 方式においては、9 ビットの I コードの各ビットは、1 又は -1 という 2 種類の符号極性を有するため、このような I コードを、クラスコードとする場合には、例えば、-1 となっているビットを 0 とみなすようにすればよい。

C E L P 方式では、コードデータに、リスト補間ビットやフレームエネルギーが

含められる場合があるが、この場合、クラスタップは、ソフト補間ピットやフレームエネルギーを用いて構成することが可能である。

特開平8-202399号公報には、合成音を、高域強調フィルタに通すことによって、その音質を改善する号法が開示されているが、本発明は、タップ係数が学習により得られる点及び用いるタップ係数が、コードによるクラス分類結果によって決まる点等において、特開平8-202339号公報に記載の発明と異なる。

次に、本発明の他の実施の形態を図面を参照して詳細に説明する。

本発明を適用した音声合成装置は、図14に示すような構成を備え、音声合成フィルタ147に与える残差信号と線形予測係数を、それぞれコード化した残差コードとAコードが多重化されたコードデータが供給されるようになっており、その残差コードとAコードから、それぞれ残差信号と線形予測係数を求め、音声合成フィルタ147に与えることで合成音が生成される。

但し、残差コードを、残差信号と残差コードとを対応付けたコードブックに基づいて、残差信号に復号した場合には、前述したように、その復号残差信号は、誤差を含むものとなり、合成音の音質が劣化する。同様に、Aコードを、線形予測係数とAコードとを対応付けたコードブックに基づいて、線形予測係数に復号した場合にも、その復号線形予測係数は、誤差を含むものとなり、合成音の音質が劣化する。

そこで、図14の音声合成装置では、学習により求めたタップ係数を用いた予測演算を行うことにより、真の残差信号と線形予測係数の予測値を求め、これらを用いることで高音質の合成音を生成する。

即ち、図14の音声合成装置では、例えば、クラス分類適応処理を利用して、復号線形予測係数が、真の線形予測係数の予測値に復号される。

クラス分類適応処理は、クラス分類処理と適応処理とからなり、クラス分類処理によって、データを、その性質に基づいてクラス分けし、各クラスごとに適応処理を施すものであり、適応処理は、前述したと同様の手法によって行われるので、ここでは、前述の説明を参照して詳細な説明は省略する。

図14の音声合成装置では、以上のようなクラス分類適応処理により、復号線

形予測係数を、真の線形予測係数（の予測値）に復号する他、復号残差信号も、真の残差信号（の予測値）に復号するようになっている。

即ち、デマルチブレクサ（DEMUX）141には、コードデータが供給されるようになっており、デマルチブレクサ141は、そこに供給されるコードデータから、フレームごとのAコードと残差コードを分離し、それを、フィルタ係数復号器142Aと残差コードブック記憶部142Eに供給する。

ここで、図14におけるコードデータに含まれるAコードと残差コードは、音声を、所定のフレームごとにLPC分析して得られる線形予測係数と残差信号を、所定のコードブックを用いて、それぞれベクトル量子化することにより得られるコードとなっている。

フィルタ係数復号器142Aは、デマルチブレクサ141から供給されるフレームごとのAコードを、そのAコードを得るときに用いられたのと同一のコードブックに基づいて、復号線形予測係数に復号し、タップ生成部143Aに供給する。

残差コードブック記憶部142Eは、デマルチブレクサ141から供給されるフレームごとの残差コードを得るときに用いられたのと同一のコードブックを記憶しており、デマルチブレクサからの残差コードを、そのコードブックに基づいて、復号残差信号に復号し、タップ生成部143Eに供給する。

タップ生成部143Aは、フィルタ係数復号器142Aから供給されるフレームごとの復号線形予測係数から、後述するクラス分類部144Aにおけるクラス分類に用いられるクラスタップとなるものと、同じく後述する予測部146における予測演算に用いられる予測タップとなるものを、それぞれ抽出する。即ち、タップ生成部143Aは、例えば、いま処理しようとしているフレームの復号線形予測係数すべてを、線形予測係数についてのクラスタップ及び予測タップとする。タップ生成部143Aは、線形予測係数についてのクラスタップをクラス分類部144Aに、予測タップを予測部146Aに、それぞれ供給する。

タップ生成部143Eは、残差コードブック記憶部142Eから供給されるフレームごとの復号残差信号から、クラスタップとなるものと、予測タップとなるものを、それぞれ抽出する。即ち、タップ生成部143Eは、例えば、いま処理

しようとしているフレームの復号残差信号のサンプル値すべてを、残差信号についてのクラスタップ及び予測タップとする。タップ生成部 143E は、残差信号についてのクラスタップをクラス分類部 144E に、予測タップを予測部 146E に、それぞれ供給する。

ここで、予測タップやクラスタップの構成パターンは、上述したパターンのものに限定されるものではない。

なお、タップ生成部 143A では、復号線形予測係数と、復号残差信号との両方の中から、線形予測係数のクラスタップや予測タップを抽出するようになるとができる。さらに、タップ生成部 143A では、A コードや残差コードからも、線形予測係数についてのクラスタップや予測タップを抽出するようになることができる。また、後段の予測部 146A や 146E が既に出力した信号や、音声合成フィルタ 147 が既に出力した合成音信号からも、線形予測係数についてのクラスタップや予測タップを抽出するようになることができる。タップ生成部 143E においても、同様にして、残差信号についてのクラスタップや予測タップを抽出することが可能である。

クラス分類部 144A は、タップ生成部 143A からの線形予測係数についてのクラスタップに基づき、注目している注目フレームである真の線形予測係数の予測値を求めようとしているフレームの線形予測係数をクラス分類し、その結果得られるクラスに対応するクラスコードを、係数メモリ 145A に出力する。

ここで、クラス分類を行う方法としては、例えば、A D R C (Adaptive Dynamic Range Coding) 等を採用することができる。

A D R C を用いる方法では、クラスタップを構成する復号線形予測係数が、A D R C 処理され、その結果得られる A D R C コードにしたがって、注目フレームの線形予測係数のクラスが決定される。

K ビット A D R C においては、例えば、クラスタップを構成する復号線形予測係数の最大値 MAX と最小値 MIN が検出され、 $DR=MAX-MIN$ を、集合の局所的なダイナミックレンジとし、このダイナミックレンジ DR に基づいて、クラスタップを構成する復号線形予測係数が K ビットに再量子化される。即ち、クラスタップを構成する復号線形予測係数の中から、最小値 MIN が減算され、その減算値が $DR/2K$ で除算

(量子化)される。そして、以上のようにして得られる、クラスタップを構成するKビットの各復号線形予測係数を、所定の順番で並べたビット列が、A D R Cコードとして出力される。従って、クラスタップが、例えば、1ビットA D R C処理された場合には、そのクラスタップを構成する各復号線形予測係数は、最小値MINが減算された後に、最大値MAXと最小値MINとの平均値で除算され、これにより、各復号線形予測係数が1ビットとされる(2値化される)。そして、その1ビットの復号線形予測係数を所定の順番で並べたビット列が、A D R Cコードとして出力される。

クラス分類部144Aには、例えば、クラスタップを構成する復号線形予測係数の値の系列を、そのままクラスコードとして出力させることも可能であるが、この場合、クラスタップが、P次の復号線形予測係数で構成され、各復号線形予測係数に、Kビットが割り当てられているとすると、クラス分類部144Aが出力するクラスコードの場合の数は、 $(2^P)^K$ 通りとなり、復号線形予測係数のビット数Kに指数的に比例した膨大な数となる。

従って、クラス分類部144Aにおいては、クラスタップの情報量を、上述のA D R C処理や、あるいはベクトル量子化等によって圧縮してから、クラス分類を行うのが好ましい。

クラス分類部144Eも、タップ生成部143Eから供給されるクラスタップに基づき、クラス分類部144Aにおける場合と同様にして、注目フレームのクラス分類を行い、その結果得られるクラスコードを、係数メモリ145Eに出力する。

係数メモリ145Aは、後述する図17の学習装置において学習処理が行われることにより得られる、クラスごとの線形予測係数についてのタップ係数を記憶しており、クラス分類部144Aが出力するクラスコードに対応するアドレスに記憶されているタップ係数を、予測部146Aに出力する。

係数メモリ145Eは、後述する図17の学習装置において学習処理が行われることにより得られる、クラスごとの残差信号についてのタップ係数を記憶しており、クラス分類部144Eが出力するクラスコードに対応するアドレスに記憶されているタップ係数を、予測部146Eに出力する。

ここで、各フレームについて、P次の線形予測係数が求められるとすると、注目フレームについて、P次の線形予測係数を、前述した式(6)の予測演算によって求めるには、Pセットのタップ係数が必要である。従って、係数メモリ145Aには、1つのクラスコードに対応するアドレスに対して、Pセットのタップ係数が記憶されている。同様の理由から、係数メモリ145Eには、各フレームにおける残差信号のサンプル点と同一数のセットのタップ係数が記憶されている。

予測部146Aは、タップ生成部143Aが出力する予測タップと、係数メモリ145Aが出力するタップ係数とを取得し、その予測タップとタップ係数とを用いて、式(6)に示した線形予測演算(積和演算)を行い、注目フレームのP次の線形予測係数(の予測値)を求めて、音声合成フィルタ147に出力する。

予測部146Eは、タップ生成部143Eが出力する予測タップと、係数メモリ145Eが出力するタップ係数とを取得し、その予測タップとタップ係数とを用いて、式(6)に示した線形予測演算を行い、注目フレームの残差信号の予測値を求めて、音声合成フィルタ147に出力する。

ここで、係数メモリ145Aは、注目フレームを構成するP次の線形予測係数の予測値それぞれを求めるためのPセットのタップ係数を出力するが、予測部146Aは、各次数の線形予測係数を、予測タップと、その次数に対応するタップ係数のセットとを用いて、式(6)の積和演算を行う。予測部146Eも同様である。

音声合成フィルタ147は、例えば、前述した図1の音声合成フィルタ29と同様に、IIR型のデジタルフィルタで、予測部146Aからの線形予測係数をIIRフィルタのタップ係数とともに、予測部146Eからの残差信号を入力信号として、その入力信号のフィルタリングを行うことにより、合成音信号を生成し、D/A変換部148に供給する。D/A変換部148は、音声合成フィルタ147からの合成音信号を、デジタル信号からアナログ信号にD/A変換し、スピーカ149に供給して出力させる。

なお、図14では、タップ生成部143Aと143Eにおいて、それぞれクラスタップを生成し、クラス分類部144Aと144Eにおいて、それぞれ、そのクラスタップに基づくクラス分類を行い、さらに、係数メモリ145Aと145E

Eから、それぞれ、そのクラス分類結果としてのクラスコードに対応する、線形予測係数と残差信号それについてのタップ係数を取得するようにしたが、線形予測係数と残差信号それについてのタップ係数は、例えば、以下のようにして取得することも可能である。

即ち、タップ生成部143Aと143E、クラス分類部144Aと144E、係数メモリ145Aと145Eを、それぞれ一体的に構成する。いま、一体的に構成したタップ生成部、クラス分類部、係数メモリを、それぞれ、タップ生成部143、クラス分類部144、係数メモリ145というものとすると、タップ生成部143には、復号線形予測係数と復号残差信号とからクラスタップを構成させ、クラス分類部144には、そのクラスタップに基づいて、クラス分類を行わせ、1つのクラスコードを出力させる。さらに、係数メモリ145には、各クラスに対応するアドレスに、線形予測係数についてのタップ係数と、残差信号についてのタップ係数との組を記憶させておき、クラス分類部144が出力するクラスコードに対応するアドレスに記憶されている線形予測係数と残差信号それについてのタップ係数の組を出力させる。そして、予測部146Aと146Eでは、このようにして、係数メモリ145から組で出力される線形予測係数についてのタップ係数と、残差信号についてのタップ係数に基づいて、それぞれ、処理を行うようにすることができる。

なお、タップ生成部143Aと143E、クラス分類部144Aと144E、係数メモリ145Aと145Eを、それぞれ別に構成する場合には、線形予測係数についてのクラス数と、残差信号についてのクラス数とは、同一になるとは限らないが、一体的に構成する場合には、線形予測係数と残差信号についてのクラス数は、同一になる。

次に、図14に示す音声合成装置を構成する音声合成フィルタ147の具体的な構成を図15に示す。

音声合成フィルタ147は、図15に示すように、P次の線形予測係数を用いるものとなっており、従って、1つの加算器151、P個の遅延回路(D)152乃至152、及びP個の乗算器153乃至153から構成されている。

乗算器153乃至153には、それぞれ、予測部146Aから供給されるP次

の線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ がセットされ、これにより、音声合成フィルタ 147 では、式(4)にしたがって演算が行われ、合成音信号が生成される。

即ち、予測部 146E が出力する残差信号 e は、加算器 151 を介して、遅延回路 152 に供給され、遅延回路 152 は、そこへの入力信号を、残差信号の 1 サンプル分だけ遅延して、後段の遅延回路 152H に出力するとともに、乗算器 153 に出力する。乗算器 153 は、遅延回路 12 の出力と、そこにセットされた線形予測係数 α_i とを乗算し、その乗算値を加算器 151 に出力する。

加算器 151 は、乗算器 153 乃至 153H の出力すべてと、残差信号 e とを加算し、その加算結果を、遅延回路 12 に供給する他、音声合成結果（合成音信号）として出力する。

次に、図 16 のフローチャートを参照して、図 14 の音声合成装置の音声合成処理について説明する。

デマルチブレクサ 141 は、そこに供給されるコードデータから、フレームごとの A コードと残差コードを順次分離し、それを、フィルタ係数復号器 142A と残差コードブック記憶部 142E に供給する。

フィルタ係数復号器 142A は、デマルチブレクサ 141 から供給されるフレームごとの A コードを、復号線形予測係数に順次復号し、タップ生成部 143A に供給し、また、残差コードブック記憶部 142E は、デマルチブレクサ 141 から供給されるフレームごとの残差コードを、復号残差信号に順次復号し、タップ生成部 143E に供給する。

タップ生成部 143A は、そこに供給される復号線形予測係数のフレームを、順次、注目フレームとし、ステップ S101において、フィルタ係数復号器 142A から供給される復号線形予測係数から、クラスタップと予測タップを生成する。さらに、ステップ S101 では、タップ生成部 143E は、残差コードブック記憶部 142E から供給される復号残差信号から、クラスタップと予測タップを生成する。タップ生成部 143A が生成したクラスタップは、クラス分類部 144A に、予測タップは、予測部 146A に、それぞれ供給され、タップ生成部 143E が生成したクラスタップは、クラス分類部 144E に、予測タップは、予測部 146E にそれぞれ供給される。

ステップS102に進み、クラス分類部144Aと144Eは、タップ生成部143Aと143Eから供給されるクラスタップに基づいて、それぞれクラス分類を行い、その結果得られるクラスコードを、係数メモリ145Aと145Eに、それぞれ供給してステップS103に進む。

ステップS103では、係数メモリ145Aと145Eは、クラス分類部144Aと144Eから供給されるクラスコードに対応するアドレスから、タップ係数を、それぞれ読み出し、予測部146Aと146Eにそれぞれ供給する。

ステップS104に進み、予測部146Aは、係数メモリ145Aが出力するタップ係数を取得し、そのタップ係数と、タップ生成部143Aからの予測タップとを用いて、式(6)に示した積和演算を行い、注目フレームの真の線形予測係数の予測値を得る。さらに、ステップS104では、予測部146Eは、係数メモリ145Eが出力するタップ係数を取得し、そのタップ係数と、タップ生成部143Eからの予測タップとを用いて、式(6)に示した積和演算を行い、注目フレームの真の残差信号(の予測値)を得る。

以上のようにして得られた残差信号及び線形予測係数は、音声合成フィルタ147に供給され、音声合成フィルタ147では、その残差信号及び線形予測係数を用いて、式(4)の演算が行われることにより、注目フレームの合成音信号が生成される。この合成音信号は、音声合成フィルタ147から、D/A変換部148を介して、スピーカ149に供給され、これにより、スピーカ149からは、その合成音信号に対応する合成音が出力される。

予測部146Aと146Eにおいて、線形予測係数と残差信号がそれ得られた後は、ステップS105に進み、まだ、注目フレームとして処理すべきフレームの復号線形予測係数及び復号残差信号があるかどうかが判定される。ステップS105において、まだ、注目フレームとして処理すべきフレームの復号線形予測係数及び復号残差信号があると判定された場合、ステップS101に戻り、次に注目フレームとすべきフレームを、新たに注目フレームとして、以下、同様の処理を繰り返す。また、ステップS105において、注目フレームとして処理すべきフレームの復号線形予測係数及び復号残差信号がないと判定された場合、音声合成処理を終了する。

図14に示す係数メモリ145A及び145Eに記憶させるタップ係数の学習処理を行う学習装置は、図17に示すような構成を備えている。

図17に示す学習装置には、学習用のデジタル音声信号が、フレーム単位で供給されるようになっており、この学習用のデジタル音声信号は、LPC分析部161A及び予測フィルタ161Eに供給される。

LPC分析部161Aは、そこに供給される音声信号のフレームを、順次、注目フレームとし、その注目フレームの音声信号をLPC分析することで、P次の線形予測係数を求める。この線形予測係数は、予測フィルタ161E及びベクトル量子化部162Aに供給されるとともに、線形予測係数についてのタップ係数を求めるための教師データとして、正規方程式加算回路166Aに供給される。

予測フィルタ161Eは、そこに供給される注目フレームの音声信号と線形予測係数を用いて、例えば、式(1)にしたがった演算を行うことにより、注目フレームの残差信号を求め、ベクトル量子化部162Eに供給するとともに、残差信号についてのタップ係数を求めるための教師データとして、正規方程式加算回路166Eに供給する。

即ち、前述した式(1)におけるsとeのZ変換を、SとEとそれぞれ表すと、式(1)は、次式のように表すことができる。

$$E = (1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots + \alpha_P z^{-P}) S \quad \dots \quad (15)$$

式(15)から、残差信号eは、音声信号sと線形予測係数 α_i との積和演算で求めることができ、従って、残差信号eを求める予測フィルタ161Eは、FIR(Finite Impulse Response)型のデジタルフィルタで構成することができる。

即ち、図18は、予測フィルタ161Eの構成例を示している。

予測フィルタ161Eには、LPC分析部161Aから、P次の線形予測係数が供給されるようになっており、従って、予測フィルタ161Eは、P個の遅延回路(D)171,乃至171, P個の乗算器172,乃至172,及び1つの加算器173から構成されている。

乗算器172,乃至172には、それぞれ、LPC分析部161Aから供給されるP次の線形予測係数のうちの $\alpha_1, \alpha_2, \dots, \alpha_P$ がセットされる。

一方、注目フレームの音声信号sは、遅延回路171と加算器173に供給さ

れる。遅延回路 171_Hは、そこへの入力信号を、残差信号の 1 サンプル分だけ遅延して、後段の遅延回路 171_Hに output するとともに、乗算器 172_Hに出力する。乗算器 172_Hは、遅延回路 171_Hの出力と、そこにセットされた線形予測係数 α _Hとを乗算し、その乗算値を加算器 173_Hに出力する。

加算器 173_Hは、乗算器 172_H乃至 172_Hの出力すべてと、音声信号 s_Hとを加算し、その加算結果を残差信号 e_Hとして出力する。

図 17 に戻り、ベクトル量子化部 162A_Hは、線形予測係数を要素とするコードベクトルとコードとを対応付けたコードブックを記憶しており、そのコードブックに基づいて、LPC 分析部 161A_Hからの注目フレームの線形予測係数で構成される特徴ベクトルをベクトル量子化し、そのベクトル量子化の結果得られる A コードを、フィルタ係数復号器 163A_Hに供給する。ベクトル量子化部 162_H信号のサンプル値を要素とするコードベクトルとコードとを対応付けたコードブックを記憶しており、そのコードブックに基づいて、予測フィルタ 161E_Hからの注目フレームの残差信号のサンプル値で構成される残差ベクトルをベクトル量子化し、そのベクトル量子化の結果得られる残差コードを、残差コードブック記憶部 163E_Hに供給する。

フィルタ係数復号器 163A_Hは、ベクトル量子化部 162A_Hが記憶しているのと同一のコードブックを記憶しており、そのコードブックに基づいて、ベクトル量子化部 162A_Hからの A コードを、復号線形予測係数に復号し、線形予測係数についてのタップ係数を求めるための生徒データとして、タップ生成部 164A_Hに供給する。ここで、図 14 のフィルタ係数復号器 142A_Hは、図 17 のフィルタ係数復号器 163A_Hと同様に構成されている。

残差コードブック記憶部 163E_Hは、ベクトル量子化部 162E_Hが記憶しているのと同一のコードブックを記憶しており、そのコードブックに基づいて、ベクトル量子化部 162E_Hからの残差コードを、復号残差信号に復号し、残差信号についてのタップ係数を求めるための生徒データとして、タップ生成部 164E_Hに供給する。ここで、図 14 の残差コードブック記憶部 142E_Hは、図 17 の残差コードブック記憶部 142E_Hと同様に構成されている。

タップ生成部 164A_Hは、図 14 のタップ生成部 143A_Hにおける場合と同様

に、フィルタ係数復号器 163A から供給される復号線形予測係数から、予測タップとクラスタップを構成し、クラスタップを、クラス分類部 165A に供給するとともに、予測タップを、正規方程式加算回路 166A に供給する。タップ生成部 164E は、図 14 のタップ生成部 143E における場合と同様に、残差コードブック記憶部 163E から供給される復号残差信号から、予測タップとクラスタップを構成し、クラスタップを、クラス分類部 165E に供給するとともに、予測タップを、正規方程式加算回路 166E に供給する。

クラス分類部 165A と 165E は、図 3 のクラス分類部 144A と 144E における場合とそれぞれ同様に、そこに供給されるクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路 166A と 166E に、それぞれ供給する。

正規方程式加算回路 166A は、LPC 分析部 161A からの教師データとしての注目フレームの線形予測係数と、タップ生成部 164A からの生徒データとしての予測タップを構成する復号線形予測係数を対象とした足し込みを行う。正規方程式加算回路 166E は、予測フィルタ 161E からの教師データとしての注目フレームの残差信号と、タップ生成部 164E からの生徒データとしての予測タップを構成する復号残差信号を対象とした足し込みを行う。

即ち、正規方程式加算回路 166A は、クラス分類部 165A から供給されるクラスコードに対応するクラスごとに、予測タップである生徒データを用い、前述した式 (13) の行列 A における各コンポーネントとなっている、生徒データどうしの乗算 ($x_{ii}x_{ii}$) と、サメーション (Σ) に相当する演算を行う。

さらに、正規方程式加算回路 166A は、やはり、クラス分類部 165A から供給されるクラスコードに対応するクラスごとに、生徒データ、即ち、予測タップを構成する復号線形予測係数及び教師データ、即ち、注目フレームの線形予測係数) を用い、式 (13) のベクトル v における各コンポーネントとなっている、生徒データと教師データの乗算 ($x_{ii}y_{ii}$) と、サメーション (Σ) に相当する演算を行う。

正規方程式加算回路 166A は、以上の足し込みを、LPC 分析部 161A から供給される線形予測係数のフレームすべてを注目フレームとして行い、これに

より、各クラスについて、線形予測係数に関する式（13）に示した正規方程式をたてる。

正規方程式加算回路166Eも、同様の足し込みを、予測フィルタ161Eから供給される残差信号のフレームすべてを注目フレームとして行い、これにより、各クラスについて、残差信号に関する式（13）に示した正規方程式をたてる。

タップ係数決定回路167Aと167Eは、正規方程式加算回路166Aと166Eにおいてクラスごとに生成された正規方程式それを解くことにより、クラスごとに、線形予測係数と残差信号についてのタップ係数をそれぞれ求め、係数メモリ168Aと168Eの、各クラスに対応するアドレスにそれぞれ供給する。

なお、学習用の音声信号として用意した音声信号によっては、正規方程式加算回路166Aや166Eにおいて、タップ係数を求めるのに必要な数の正規方程式が得られないクラスが生じる場合があり得るが、タップ係数決定回路167Aと167Eは、そのようなクラスについては、例えば、デフォルトのタップ係数を出力する。

係数メモリ168Aと168Eは、タップ係数決定回路167Aと167Eからそれぞれ供給されるクラスごとの線形予測係数と残差信号についてのタップ係数を、それぞれ記憶する。

次に、図19に示すフローチャートを参照して、図17の学習装置の学習処理について説明する。

学習装置には、学習用の音声信号が供給され、ステップS111では、その学習用の音声信号から、教師データと生徒データが生成される。

即ち、LPC分析部161Aは、学習用の音声信号のフレームを、順次、注目フレームとし、その注目フレームの音声信号をLPC分析することで、P次の線形予測係数を求め、教師データとして、正規方程式加算回路166Aに供給する。さらに、この線形予測係数は、予測フィルタ161E及びベクトル量子化部162Aにも供給され、ベクトル量子化部162Aは、LPC分析部161Aからの注目フレームの線形予測係数で構成される特徴ベクトルをベクトル量子化し、そのベクトル量子化の結果得られるAコードを、フィルタ係数復号器163Aに供給

する。フィルタ係数復号器 163A は、ベクトル量子化部 162A からの A コードを、復号線形予測係数に復号し、その復号線形予測係数を、生徒データとして、タップ生成部 164A に供給する。

一方、注目フレームの線形予測係数を、LPC 分析部 161A から受信した予測フィルタ 161E は、その線形予測係数と、注目フレームの学習用の音声信号とを用いて、前述した式(1)にしたがった演算を行うことにより、注目フレームの残差信号を求め、教師データとして、正規方程式加算回路 166E に供給する。この残差信号は、ベクトル量子化部 162E にも供給され、ベクトル量子化部 162E は、予測フィルタ 161E からの注目フレームの残差信号のサンプル値で構成される残差ベクトルをベクトル量子化し、そのベクトル量子化の結果得られる残差コードを、残差コードブック記憶部 163E に供給する。残差コードブック記憶部 163E は、ベクトル量子化部 162E からの残差コードを、復号残差信号に復号し、その復号残差信号を、生徒データとして、タップ生成部 164E に供給する。

そして、ステップ S112 に進み、タップ生成部 164A が、フィルタ係数復号器 163A から供給される復号線形予測係数から、線形予測係数についての予測タップとクラスタップを構成するとともに、タップ生成部 164E が、残差コードブック記憶部 163E から供給される復号残差信号から、残差信号についての予測タップとクラスタップを構成する。線形予測係数についてのクラスタップは、クラス分類部 165A に供給され、予測タップは、正規方程式加算回路 166A に供給される。また、残差信号についてのクラスタップは、クラス分類部 165E に供給され、予測タップは、正規方程式加算回路 166E に供給される。

その後、ステップ S113において、クラス分類部 165A が、線形予測係数についてのクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路 166A に供給するとともに、クラス分類部 165E が、残差信号についてのクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路 166E に供給する。

ステップ S114 に進み、正規方程式加算回路 166A は、LPC 分析部 161A からの教師データとしての注目フレームの線形予測係数、及びタップ生成部

164Aからの生徒データとしての予測タップを構成する復号線形予測係数を対象として、式(13)の行列Aとベクトルvの、上述したような足し込みを行う。さらに、ステップS114では、正規方程式加算回路166Eが、予測フィルタ161Eからの教師データとしての注目フレームの残差信号、及びタップ生成部164Eからの生徒データとしての予測タップを構成する復号残差信号を対象として、式(13)の行列Aとベクトルvの、上述したような足し込みを行い、ステップS115に進む。

ステップS115では、まだ、注目フレームとして処理すべきフレームの学習用の音声信号があるかどうかが判定される。ステップS115において、まだ、注目フレームとして処理すべきフレームの学習用の音声信号があると判定された場合、ステップS111に戻り、次のフレームを新たに注目フレームとして、以下、同様の処理が繰り返される。

ステップS105において、注目フレームとして処理すべきフレームの学習用の音声信号がないと判定された場合、即ち、正規方程式加算回路166Aと166Eにおいて、各クラスについて、正規方程式が得られた場合、ステップS116に進み、タップ係数決定回路167Aは、各クラスごとに生成された正規方程式を解くことにより、各クラスごとに、線形予測係数についてのタップ係数を求め、係数メモリ168Aの、各クラスに対応するアドレスに供給して記憶させる。さらに、タップ係数決定回路167Eも、各クラスごとに生成された正規方程式を解くことにより、各クラスごとに、残差信号についてのタップ係数を求め、係数メモリ168Eの、各クラスに対応するアドレスに供給して記憶させ、処理を終了する。

以上のようにして、係数メモリ168Aに記憶された各クラスごとの線形予測係数についてのタップ係数が、図14の係数メモリ145Aに記憶されているとともに、係数メモリ168Eに記憶された各クラスごとの残差信号についてのタップ係数が、図14の係数メモリ145Eに記憶されている。

従って、図14の係数メモリ45Aに記憶されたタップ係数は、線形予測演算を行うことにより得られる真の線形予測係数の予測値の予測誤差（ここでは、自乗誤差）が、統計的に最小になるように学習を行うことにより求められたもので

あり、また、係数メモリ 145E に記憶されたタップ係数も、線形予測演算を行うことにより得られる真の残差信号の予測値の予測誤差（自乗誤差）が、統計的に最小になるように学習を行うことにより求められたものであるから、図 14 の予測部 146A と 146E が出力する線形予測係数と残差信号は、それぞれ真の線形予測係数と残差信号にほぼ一致することとなり、その結果、これらの線形予測係数と残差信号によって生成される合成音は、歪みの少ない、高音質のものとなる。

なお、図 14 に示す音声合成装置において、上述したように、例えば、タップ生成部 143A に、復号線形予測係数と復号残差信号との両方から、線形予測係数のクラスタップや予測タップを抽出させるようにする場合には、図 17 のタップ生成部 164A にも、復号線形予測係数と復号残差信号との両方から、線形予測係数のクラスタップや予測タップを抽出させるようにする必要がある。タップ生成部 164E についても同様である。

また、図 14 に示す 3 の音声合成装置において、上述したように、タップ生成部 143A と 143E、クラス分類部 144A と 144E、係数メモリ 145A と 145E を、それぞれ一体的に構成する場合には、図 17 に示す学習装置においても、タップ生成部 164A と 164E、クラス分類部 165A と 165E、正規方程式加算回路 166A と 166E、タップ係数決定回路 167A と 167E、係数メモリ 168A と 168E を、それぞれ一体的に構成する必要がある。この場合、正規方程式加算回路 166A と 166E を一体的に構成した正規方程式加算回路では、LPC 分析部 161A が出力する線形予測係数と、予測フィルタ 161E が出力する残差信号との両方を、一度に、教師データとともに、フィルタ係数復号器 163A が出力する復号線形予測係数と、残差コードブック記憶部 163E が出力する復号残差信号との両方を、一度に、生徒データとして、正規方程式がたてられ、タップ係数決定回路 167A と 167E とを一体的に構成したタップ係数決定回路では、その正規方程式を解くことにより、クラスごとの、線形予測係数と残差信号それぞれについてのタップ係数が、一度に求められる。

次に、本発明を適用した伝送システムの一例を、図 20 を参照して説明する。

ここで、システムとは、複数の装置が論理的に集合した物をいい、各構成の装置が同一筐体中にあるか否かは問わない。

この伝送システムでは、携帯電話機 181_1 と 181_2 が、基地局 182_1 と 182_2 、それぞれの間で無線による通信を行うとともに、基地局 182_1 と 182_2 、それが交換局 83 との間で通信を行うことにより、最終的には、携帯電話機 181_1 と 181_2 との間において、基地局 182_1 と 182_2 並びに交換局 183 を通して音声の送受信を行うことができるようになっている。なお、基地局 182_1 と 182_2 は、同一の基地局であってもよいし異なる基地局であってもよい。

ここで、以下、特に区別する必要がない限り、携帯電話機 181_1 と 181_2 を携帯電話機 181 と記述する。

図 21 は、図 20 に示す携帯電話機 181 の構成例を示している。

アンテナ 191 は、基地局 182_1 と 182_2 からの電波を受信し、その受信信号を、変復調部 192 に供給するとともに、変復調部 192 からの信号を電波で基地局 182_1 又は 182_2 に送信する。変復調部 192 は、アンテナ 191 からの信号を復調し、その結果得られる、前述した図 1 で説明したようなコードデータを、受信部 194 に供給する。変復調部 192 は、送信部 193 から供給される、図 1 で説明したようなコードデータを変調し、その結果得られる変調信号を、アンテナ 191 に供給する。送信部 193 は、図 1 に示した送信部と同様に構成され、そこに入力されるユーザの音声を、コードデータに符号化して、変復調部 192 に供給する。受信部 194 は、変復調部 192 からのコードデータを受信し、そのコードデータから、図 14 の音声合成装置における場合と同様の高音質の音声を復号して出力する。

即ち、図 21 に示す受信部 194 は、図 22 に示すような構成を備えている。の構成例を示している。なお、図中、図 2 における場合と対応する部分については、同一の符号を付してあり、以下ではその説明は適宜省略する。

タップ生成部 101 には、チャネルデコーダ 21 が出力する、フレーム又はサブフレームごとのLコード、Gコード、Iコード、及びAコードが供給されるようになっており、タップ生成部 101 は、そのLコード、Gコード、Iコード、及びAコードから、クラスタップとするものを抽出し、クラス分類部 104 に供

給する。ここで、タップ生成部 101 が生成するような、レコード等で構成されるクラスタップを、以下、適宜、第1のクラスタップという。

タップ生成部 102 には、演算器 28 が出力する、フレーム又はサブフレームごとの残差信号 e が供給されるようになっており、タップ生成部 102 は、その残差信号から、クラスタップとするもの（サンプル点）を抽出し、クラス分類部 104 に供給する。さらに、タップ生成部 102 は、演算器 28 からの残差信号から、予測タップとするものを抽出し、予測部 106 に供給する。ここで、タップ生成部 102 が生成するような、残差信号で構成されるクラスタップを、以下、適宜、第2のクラスタップという。

タップ生成部 103 には、フィルタ係数復号器 25 が出力する、フレームごとの線形予測係数 α , が供給されるようになっており、タップ生成部 103 は、その線形予測係数から、クラスタップとするものを抽出し、クラス分類部 104 に供給する。さらに、タップ生成部 103 は、フィルタ係数復号器 25 からの線形予測係数から、予測タップとするものを抽出し、予測部 107 に供給する。ここで、タップ生成部 103 が生成するような、線形予測係数で構成されるクラスタップを、以下、適宜、第3のクラスタップという。

クラス分類部 104 は、タップ生成部 101 乃至 103 それから供給される第1乃至第3のクラスタップをまとめて、最終的なクラスタップとし、その最終的なクラスタップに基づいて、クラス分類を行い、そのクラス分類結果としてのクラスコードを、係数メモリ 105 に供給する。

係数メモリ 105 は、後述する図 23 の学習装置において学習処理が行われることにより得られる、クラスごとの線形予測係数についてのタップ係数と、残差信号についてのタップ係数を記憶しており、クラス分類部 104 が出力するクラスコードに対応するアドレスに記憶されているタップ係数を、予測部 106 と 107 に供給する。なお、係数メモリ 105 から予測部 106 に対しては、残差信号についてのタップ係数 W_e が供給され、係数メモリ 105 から予測部 107 に対しては、線形予測係数についてのタップ係数 W_a が供給される。

予測部 106 は、図 14 の予測部 146E と同様に、タップ生成部 102 が出力する予測タップと、係数メモリ 105 が出力する残差信号についてのタップ係

数とを取得し、その予測タップとタップ係数とを用いて、式(6)に示した線形予測演算を行う。これにより、予測部106は、注目フレームの残差信号の予測値 e_m を求めて、音声合成フィルタ29に、入力信号として供給する。

予測部107は、図14の予測部146Aと同様に、タップ生成部103が出力する予測タップと、係数メモリ105が出力する線形予測係数についてのタップ係数とを取得し、その予測タップとタップ係数とを用いて、式(6)に示した線形予測演算を行う。これにより、予測部107は、注目フレームの線形予測係数の予測値 $m\alpha$ を求めて音声合成フィルタ29に供給する。

以上のように構成される受信部194では、基本的には、図16に示したフローチャートにしたがった処理と同様の処理が行われることで、高音質の合成音が、音声の復号結果として出力される。

即ち、チャネルデコーダ21は、そこに供給されるコードデータから、Lコード、Gコード、Iコード、Aコードを分離し、それぞれを、適応コードブック記憶部22、ゲイン復号器23、励起コードブック記憶部24、フィルタ係数復号器25に供給する。さらに、Lコード、Gコード、Iコード、及びAコードは、タップ生成部101にも供給される。

適応コードブック記憶部22、ゲイン復号器23、励起コードブック記憶部24、演算器26乃至28では、前述した図1の適応コードブック記憶部9、ゲイン復号器10、励起コードブック記憶部11、演算器12乃至14における場合と同様の処理が行われ、これにより、Lコード、Gコード、及びIコードが、残差信号 e に復号される。この復号残差信号は、演算器28からタップ生成部102に供給される。

フィルタ係数復号器25は、図1で説明したように、そこに供給されるAコードを、復号線形予測係数に復号し、タップ生成部103に供給する。

タップ生成部101は、そこに供給されるLコード、Gコード、Iコード、及びAコードのフレームを、順次、注目フレームとし、ステップS101(図16参照)において、チャネルデコーダ21からのLコード、Gコード、Iコード、及びAコードから、第1のクラスタップを生成し、クラス分類部104に供給する。ステップS101では、タップ生成部102が、演算器28からの復号残差

信号から、第2のクラスタップを生成し、クラス分類部104に供給するとともに、タップ生成部103が、フィルタ係数復号器25からの線形予測係数から、第3のクラスタップを生成し、クラス分類部104に供給する。また、ステップS101では、タップ生成部102が、演算器28からの残差信号から、予測タップとするものを抽出し、予測部106に供給するとともに、タップ生成部103が、フィルタ係数復号器25からの線形予測係数から、予測タップを生成し、予測部107に供給する。

ステップS102に進み、クラス分類部104は、タップ生成部101乃至103それそれから供給される第1乃至第3のクラスタップをまとめた、最終的なクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、係数メモリ105に供給してステップS103に進む。

ステップS103では、係数メモリ105は、クラス分類部104から供給されるクラスコードに対応するアドレスから、残差信号と線形予測係数それについてのタップ係数を読み出し、残差信号についてのタップ係数を、予測部106に供給するとともに、線形予測係数についてのタップ係数を、予測部107に供給する。

ステップS104に進み、予測部106は、係数メモリ105が出力する残差信号についてのタップ係数を取得し、そのタップ係数と、タップ生成部102からの予測タップとを用いて、式(6)に示した積和演算を行い、注目フレームの真の残差信号の予測値を得る。さらに、ステップS104では、予測部107は、係数メモリ105が出力する線形予測係数についてのタップ係数を取得し、そのタップ係数と、タップ生成部103からの予測タップとを用いて、式(6)に示した積和演算を行い、注目フレームの真の線形予測係数の予測値を得る。

以上のようにして得られた残差信号及び線形予測係数は、音声合成フィルタ29に供給され、音声合成フィルタ29では、その残差信号及び線形予測係数を用いて、式(4)の演算が行われることにより、注目フレームの合成音信号が生成される。この合成音信号は、音声合成フィルタ29から、D/A変換部30を介して、スピーカ31に供給され、これにより、スピーカ31からは、その合成音信号に対応する合成音が出力される。

予測部 106 と 107において、残差信号と線形予測係数がそれぞれ得られた後は、ステップ S105 に進み、まだ、注目フレームとして処理すべきフレームの L コード、G コード、I コード、及び A コードがあるかどうかが判定される。ステップ S105 において、まだ、注目フレームとして処理すべきフレームの L コード、G コード、I コード、及び A コードがあると判定された場合、ステップ S101 に戻り、次に注目フレームとすべきフレームを、新たに注目フレームとして、以下、同様の処理を繰り返す。また、ステップ S105 において、注目フレームとして処理すべきフレームの L コード、G コード、I コード、及び A コードがないと判定された場合処理を終了する。

次に、図 22 に示す係数メモリ 105 に記憶させるタップ係数の学習処理を行う学習装置の一例を図 23 を参照して説明する。なお、以下の説明では、図 12 に示した学習装置と共に通する部分には共通する符号を付す。

マイクロフォン 201 乃至コード決定部 215 は、図 1 のマイクロフォン 1 乃至コード決定部 15 とそれ同様に構成される。そして、マイクロフォン 201 には、学習用の音声信号が入力されるようになっており、従って、マイクロフォン 201 乃至コード決定部 215 では、その学習用の音声信号に対して、図 1 における場合と同様の処理が施される。

予測フィルタ 111E には、A/D 変換部 202 が出力する、デジタル信号とされた学習用の音声信号と、LPC 分析部 204 が出力する線形予測係数が供給される。また、タップ生成部 112A には、ベクトル量子化部 205 が出力する線形予測係数、即ち、ベクトル量子化に用いられるコードブックのコードベクトル（セントロイドベクトル）を構成する線形予測係数が供給され、タップ生成部 112E には、演算器 214 が出力する残差信号、すなわち、音声合成フィルタ 206 に供給されるのと同一の残差信号が供給される。さらに、正規方程式加算回路 114A には、LPC 分析部 204 が出力する線形予測係数が供給され、タップ生成部 117 には、コード決定部 215 が出力する L コード、G コード、I コード、及び A コードが供給される。

予測フィルタ 111E は、A/D 変換部 202 から供給される学習用の音声信号のフレームを、順次、注目フレームとして、その注目フレームの音声信号と、

LPC分析部204から供給される線形予測係数を用いて、例えば、式(1)にしたがった演算を行うことにより、注目フレームの残差信号を求める。この残差信号は、教師データとして、正規方程式加算回路114Eに供給される。

タップ生成部112Aは、ベクトル量子化部205から供給される線形予測係数から、図11のタップ生成部103における場合と同一の予測タップと第3のクラスタップを構成し、第3のクラスタップを、クラス分類部113A及び113Eに供給するとともに、予測タップを、正規方程式加算回路114Aに供給する。

タップ生成部112Eは、演算器214から供給される残差信号から、図22のタップ生成部102における場合と同一の予測タップと第2のクラスタップを構成し、第2のクラスタップを、クラス分類部113A及び113Eに供給するとともに、予測タップを、正規方程式加算回路114Eに供給する。

クラス分類部113A及び113Eには、タップ生成部112Aと112Eから、それぞれ第3と第2のクラスタップが供給される他、タップ生成部117から第1のクラスタップも供給される。そして、クラス分類部113Aと113Eは、図22のクラス分類部104における場合と同様に、そこに供給される第1乃至第3のクラスタップをまとめて、最終的なクラスタップとし、その最終的なクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路114Aと114Eにそれぞれ供給する。

正規方程式加算回路114Aは、LPC分析部204からの注目フレームの線形予測係数を、教師データとして受信するとともに、タップ生成部112Aからの予測タップを、生徒データとして受信し、その教師データ及び生徒データを対象として、クラス分類部113Aからのクラスコードごとに、図17の正規方程式加算回路166Aにおける場合と同様の足し込みを行うことにより、各クラスについて、線形予測係数に関する式(13)に示した正規方程式をたてる。正規方程式加算回路114Eは、予測フィルタ111Eからの注目フレームの残差信号を、教師データとして受信するとともに、タップ生成部112Eからの予測タップを、生徒データとして受信し、その教師データ及び生徒データを対象として、クラス分類部113Eからのクラスコードごとに、図17の正規方程式加算回路

166Eにおける場合と同様の足し込みを行うことにより、各クラスについて、残差信号に関する式(13)に示した正規方程式をたてる。

タップ係数決定回路115Aと115Eは、正規方程式加算回路114Aと114Eにおいてクラスごとに生成された正規方程式それぞれを解くことにより、クラスごとに、線形予測係数と残差信号についてのタップ係数をそれぞれ求め、係数メモリ116Aと116Eの、各クラスに対応するアドレスにそれぞれ供給する。

なお、学習用の音声信号として用意する音声信号によっては、正規方程式加算回路114Aや114Eにおいて、タップ係数を求めるのに必要な数の正規方程式が得られないクラスが生じる場合があり得るが、タップ係数決定回路115Aと115Eは、そのようなクラスについては、例えば、デフォルトのタップ係数を出力する。

係数メモリ116Aと116Eは、タップ係数決定回路115Aと115Eから、それぞれ供給されるクラスごとの線形予測係数と残差信号についてのタップ係数を、それぞれ記憶する。

タップ生成部117は、コード決定部215から供給されるLコード、Gコード、Iコード、及びAコードから、図22のタップ生成部101における場合と同一の第1のクラスタップを生成し、クラス分類部113A及び113Eに供給する。

以上のように構成される学習装置では、基本的には、図19に示したフローチャートにしたがった処理と同様の処理が行われることで、高音質の合成音を得るためにタップ係数が求められる。

学習装置には、学習用の音声信号が供給され、ステップS111において、その学習用の音声信号から、教師データと生徒データが生成される。

即ち、学習用の音声信号は、マイクロフォン201に入力され、マイクロフォン201乃至コード決定部215は、図1のマイクロフォン1乃至コード決定部15における場合とそれ同様の処理を行う。

その結果、LPC分析部204で得られる線形予測係数は、教師データとして、正規方程式加算回路114Aに供給される。また、この線形予測係数は、予測フ

ィルタ 111E にも供給される。さらに、演算器 214 で得られる残差信号は、生徒データとして、タップ生成部 112E に供給される。

A/D 変換部 202 が output するデジタルの音声信号は、予測フィルタ 111E に供給され、ベクトル量子化部 205 が output する線形予測係数は、生徒データとして、タップ生成部 112A に供給される。さらに、コード決定部 215 が output する L コード、G コード、I コード、及び A コードは、タップ生成部 117 に供給される。

予測フィルタ 111E は、A/D 変換部 202 から供給される学習用の音声信号のフレームを、順次、注目フレームとして、その注目フレームの音声信号と、LPC 分析部 204 から供給される線形予測係数を用いて、式(1)にしたがった演算を行うことにより、注目フレームの残差信号を求める。この予測フィルタ 111E で得られる残差信号は、教師データとして、正規方程式加算回路 114E に供給される。

以上のようにして、教師データと生徒データが得られた後は、ステップ S112 に進み、タップ生成部 112A が、ベクトル量子化部 205 から供給される線形予測係数から、線形予測係数についての予測タップと第 3 のクラスタップを生成するとともに、タップ生成部 112E が、演算器 214 から供給される残差信号から、残差信号についての予測タップと第 2 のクラスタップを生成する。さらに、ステップ S112 では、タップ生成部 117 が、コード決定部 215 から供給される L コード、G コード、I コード、及び A コードから、第 1 のクラスタップを生成する。

線形予測係数についての予測タップは、正規方程式加算回路 114A に供給され、残差信号についての予測タップは、正規方程式加算回路 114E に供給される。また、第 1 乃至第 3 のクラスタップは、クラス分類回路 113A 及び 113E に供給される。

その後、ステップ S113において、クラス分類部 113A と 113E が、第 1 乃至第 3 のクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路 114A と 114E に、それぞれ供給する。

ステップ S114 に進み、正規方程式加算回路 114A は、LPC 分析部 20

4からの教師データとしての注目フレームの線形予測係数、及びタップ生成部112Aからの生徒データとしての予測タップを対象として、式(13)の行列Aとベクトルvの、上述したような足し込みを、クラス分類部113Aからのクラスコードごとに行う。さらに、ステップS114では、正規方程式加算回路114Eが、予測フィルタ111Eからの教師データとしての注目フレームの残差信号、及びタップ生成部112Eからの生徒データとしての予測タップを対象として、式(13)の行列Aとベクトルvの、上述したような足し込みを、クラス分類部113Eからのクラスコードごとに行い、ステップS115に進む。

ステップS115では、まだ、注目フレームとして処理すべきフレームの学習用の音声信号があるかどうかが判定される。ステップS115において、まだ、注目フレームとして処理すべきフレームの学習用の音声信号があると判定された場合、ステップS111に戻り、次のフレームを新たに注目フレームとして、以下同様の処理が繰り返される。

ステップS115において、注目フレームとして処理すべきフレームの学習用の音声信号がないと判定された場合、即ち、正規方程式加算回路114Aと114Eそれぞれにおいて、各クラスについて、正規方程式が得られた場合、ステップS116に進み、タップ係数決定回路115Aは、各クラスごとに生成された正規方程式を解くことにより、各クラスごとに、線形予測係数についてのタップ係数を求め、係数メモリ116Aの、各クラスに対応するアドレスに供給して記憶させる。さらに、タップ係数決定回路115Eも、各クラスごとに生成された正規方程式を解くことにより、各クラスごとに、残差信号についてのタップ係数を求め、係数メモリ116Eの、各クラスに対応するアドレスに供給して記憶させ処理を終了する。

以上のようにして、係数メモリ116Aに記憶された各クラスごとの線形予測係数についてのタップ係数と、係数メモリ116Eに記憶された各クラスごとの残差信号についてのタップ係数が、図22の係数メモリ105に記憶されている。

従って、図22の係数メモリ105に記憶されたタップ係数は、線形予測演算を行うことにより得られる真の線形予測係数や残差信号の予測値の予測誤差(自乗誤差)が、統計的に最小になるように学習を行うことにより求められたもので

あるから、図22の予測部106と107が出力する残差信号と線形予測係数は、それぞれ真の残差信号と線形予測係数にほぼ一致することとなり、その結果、これらの残差信号と線形予測係数によって生成される合成音は、歪みの少ない、高音質のものとなる。

上述した一連の処理は、ハードウェアにより行うこともできるし、ソフトウェアにより行うこともできる。一連の処理をソフトウェアによって行う場合には、そのソフトウェアを構成するプログラムが、汎用のコンピュータ等にインストールされる。

上述した一連の処理を実行するプログラムがインストールされるコンピュータは、前述した図13に示すように構成され、図13に示したコンピュータと同様の動作が実行されるので、その詳細な説明は省略する。

次に、本発明の更に他の実施の形態を図面を参照して詳細に説明する。

この音声合成装置には、音声合成フィルタ244に与える残差信号と線形予測係数を、それぞれベクトル量子化等によってコード化した残差コードとAコードが多重化されたコードデータが供給されるようになっており、その残差コードとAコードから、それぞれ残差信号と線形予測係数を復号し、音声合成フィルタ244に与えることで、合成音が生成されるようになっている。さらに、この音声合成装置では、音声合成フィルタ244で生成された合成音と、学習により求めたタップ係数を用いた予測演算を行うことにより、その合成音の音質を向上させた高音質の音声（合成音）を求めて出力するようになっている。

即ち、図24に示す音声合成装置では、例えば、クラス分類適応処理を利用して、合成音が、真の高音質の音声の予測値に復号される。

クラス分類適応処理は、クラス分類処理と適応処理とからなり、クラス分類処理によって、データを、その性質に基づいてクラス分けし、各クラスごとに適応処理を施すものであり、適応処理は、前述したと同様の手法によって行われるので、ここでは、前述の説明を参照して詳細な説明は省略する。

図24に示す音声合成装置では、以上のようなクラス分類適応処理により、復号線形予測係数を、真の線形予測係数（の予測値）に復号する他、復号残差信号も、真の残差信号（の予測値）に復号するようになっている。

即ち、デマルチブレクサ（DEMUX）241には、コードデータが供給されるようになっており、デマルチブレクサ241は、そこに供給されるコードデータから、フレームごとのAコードと残差コードを分離する。そして、デマルチブレクサは、Aコードを、フィルタ係数復号器242、並びにタップ生成部245及び246に供給し、残差コードを、残差コードブック記憶部243、並びにタップ生成部245及び246に供給する。

ここで、図24におけるコードデータに含まれるAコードと残差コードは、音声をLPC分析して得られる線形予測係数と残差信号を、所定のコードブックを用いて、それぞれベクトル量子化することにより得られるコードとなっている。

フィルタ係数復号器242は、デマルチブレクサ241から供給されるフレームごとのAコードを、そのAコードを得るときに用いられたのと同一のコードブックに基づいて、線形予測係数に復号し、音声合成フィルタ244に供給する。

残差コードブック記憶部243は、デマルチブレクサ241から供給されるフレームごとの残差コードを、その残差コードを得るときに用いられたのと同一のコードブックに基づいて、残差信号に復号し、音声合成フィルタ244に供給する。

音声合成フィルタ244は、例えば、前述した図2の音声合成フィルタ29と同様に、IIR型のデジタルフィルタで、フィルタ係数復号器242からの線形予測係数をIIRフィルタのタップ係数とともに、残差コードブック記憶部243からの残差信号を入力信号として、その入力信号のフィルタリングを行うことにより、合成音を生成し、タップ生成部245及び246に供給する。

タップ生成部245は、音声合成フィルタ244から供給される合成音のサンプル値、並びにデマルチブレクサ241から供給される残差コード及びAコードから、後述する予測部249における予測演算に用いられる予測タップとなるものを抽出する。即ち、タップ生成部245は、例えば、高音質の音声の予測値を求めるとしているフレームである注目フレームの合成音のサンプル値、残差コード、及びAコードすべてを、予測タップとする。そして、タップ生成部245は、予測タップを、予測部249に供給する。

タップ生成部246は、音声合成フィルタ244から供給される合成音のサン

フル値、並びにデマルチフレクサ 241 から供給されるフレーム又はサブフレームごとのAコード及び残差コードから、クラスタップとなるものを抽出する。即ち、タップ生成部 246 は、例えば、タップ生成部 246 と同様に、注目フレームの合成音のサンプル値、並びにAコード及び残差コードすべてを、クラスタップとする。そして、タップ生成部 246 は、クラスタップを、クラス分類部 247 に供給する。

ここで、予測タップやクラスタップの構成パターンは、上述したパターンのものに限定されるものではない。また、上述の場合には、同一のクラスタップ及び予測タップを構成するようにしたが、クラスタップと予測タップとは、異なる構成とすることが可能である。

さらに、タップ生成部 245 や 246 では、図 24 において点線で示すように、フィルタ係数復号器 242 が出力する、Aコードから得られる線形予測係数や、残差コードブック記憶部 243 が出力する、残差コードから得られる残差信号等の中からも、クラスタップや予測タップを抽出するようにすることができる。

クラス分類部 247 は、タップ生成部 246 からのクラスタップに基づき、注目している注目フレームの音声のサンプル値についてクラス分類を行い、その結果得られるクラスに対応するクラスコードを、係数メモリ 248 に出力する。

ここで、クラス分類部 247 には、例えば、クラスタップとしての注目フレームの合成音のサンプル値、並びにAコード及び残差コードを構成するビットの系列そのものをクラスコードとして出力させることができる。

係数メモリ 248 は、後述する図 27 の学習装置において学習処理が行われることにより得られる、クラスごとのタップ係数を記憶しており、クラス分類部 247 が出力するクラスコードに対応するアドレスに記憶されているタップ係数を、予測部 249 に出力する。

ここで、各フレームについて、Nサンプルの高音質の音声が求められるとすると、注目フレームについて、Nサンプルの音声を、式(6)の予測演算によって求めるには、Nセットのタップ係数が必要である。従って、この場合は、係数メモリ 248 には、1つのクラスコードに対応するアドレスに対して、Nセットのタップ係数が記憶されている。

予測部 249 は、タップ生成部 245 が出力する予測タップと、係数メモリ 248 が出力するタップ係数とを取得し、その予測タップとタップ係数とを用いて、前述した式（6）に示した線形予測演算（積和演算）を行い、注目フレームの高音質の音声の予測値を求めて、D/A 変換部 250 に出力する。

ここで、係数メモリ 248 は、上述したように、注目フレームの音声の Nサンプルそれぞれを求めるための N セットのタップ係数を出力するが、予測部 249 は、各サンプル値について、予測タップと、そのサンプル値に対応するタップ係数のセットとを用い、式（6）の積和演算を行う。

D/A 変換部 250 は、予測部 249 からの音声の予測値を、ディジタル信号からアナログ信号に D/A 変換し、スピーカ 51 に供給して出力させる。

次に、図 4 は、図 24 に示した音声合成フィルタ 244 の具体的な構成を図 25 に示す。図 25 に示す音声合成フィルタ 244 は、P 次の線形予測係数を用いるものとなっており、従って、1 つの加算器 261、P 個の遅延回路（D）261 乃至 262、及び P 個の乗算器 263 乃至 263_P から構成されている。

乗算器 263 乃至 263_P には、それぞれ、フィルタ係数復号器 242 から供給される P 次の線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_P$ がセットされ、これにより、音声合成フィルタ 244 では、式（4）にしたがって演算が行われ、合成音が生成される。

即ち、残差コードブック記憶部 243 が出力する残差信号 e は、加算器 261 を介して、遅延回路 262₁ に供給され、遅延回路 262₁ は、そこへの入力信号を、残差信号の 1 サンプル分だけ遅延して、後段の遅延回路 262₂ に output するとともに、演算器 263₁ に output する。乗算器 263₁ は、遅延回路 262₁ の出力と、そこにセットされた線形予測係数 α_1 とを乗算し、その乗算値を、加算器 261 に output する。

加算器 261 は、乗算器 263 乃至 263_P の出力すべてと、残差信号 e とを加算し、その加算結果を、遅延回路 262₁ に供給する他、音声合成結果（合成音）として出力する。

次に、図 26 のフローチャートを参照して、図 24 の音声合成装置の音声合成処理について説明する。

デマルチブレクサ241は、そこに供給されるコードデータから、フレームごとのAコードと残差コードを順次分離し、それぞれを、フィルタ係数復号器242と残差コードブック記憶部243に供給する。さらに、デマルチブレクサ241は、Aコード及び残差コードを、タップ生成部245及び246にも供給する。

フィルタ係数復号器242は、デマルチブレクサ241から供給されるフレームごとのAコードを、線形予測係数に順次復号し、音声合成フィルタ244に供給する。また、残差コードブック記憶部243は、デマルチブレクサ241から供給されるフレームごとの残差コードを、残差信号に順次復号し、音声合成フィルタ244に供給する。

音声合成フィルタ244では、そこに供給される残差信号及び線形予測係数を用いて、式(4)の演算が行われることにより、注目フレームの合成音が生成される。この合成音は、タップ生成部245及び46に供給される。

タップ生成部245は、そこに供給される合成音のフレームを、順次、注目フレームとし、ステップS201において、音声合成フィルタ244から供給される合成音のサンプル値、並びにデマルチブレクサ241から供給されるAコード及び残差コードから、予測タップを生成し、予測部249に出力する。さらに、ステップS201では、タップ生成部246が、音声合成フィルタ244から供給される合成音、並びにデマルチブレクサ241から供給されるAコード及び残差コードから、クラスタップを生成しクラス分類部247に出力する。

そして、ステップS202に進み、クラス分類部247は、タップ生成部246から供給されるクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、係数メモリ248に供給して、ステップS203に進む。

ステップS203では、係数メモリ248は、クラス分類部247から供給されるクラスコードに対応するアドレスから、タップ係数を読み出し、予測部249に供給する。

そして、ステップS204に進み、予測部249は、係数メモリ248が出力するタップ係数を取得し、そのタップ係数と、タップ生成部245からの予測タップとを用いて、式(6)に示した積和演算を行い、注目フレームの高音質の音声の予測値を得る。この高音質の音声は、予測部249からD/A変換部250

を介して、スピーカ 251 に供給されて出力される。

予測部 249において、注目フレームの高音質の音声が得られた後は、ステップ S205に進み、まだ、注目フレームとして処理すべきフレームがあるかどうかが判定される。ステップ S205において、まだ、注目フレームとして処理すべきフレームがあると判定された場合、ステップ S201に戻り、次に注目フレームとすべきフレームを、新たに注目フレームとして、以下、同様の処理を繰り返す。また、ステップ S205において、注目フレームとして処理すべきフレームがないと判定された場合、音声合成処理を終了する。

次に、図 27 は、図 24 に示す係数メモリ 248 に記憶させるタップ係数の学習処理を行う学習装置の一例を示すブロック図である。

図 27 に示す学習装置には、学習用の高音質のデジタル音声信号が、所定のフレーム単位で供給されるようになっており、この学習用のデジタル音声信号は、LPC 分析部 271 及び予測フィルタ 274 に供給される。さらに、学習用のデジタル音声信号は、教師データとして、正規方程式加算回路 281 にも供給される。

LPC 分析部 271 は、そこに供給される音声信号のフレームを、順次、注目フレームとし、その注目フレームの音声信号を LPC 分析することで、P 次の線形予測係数を求め、ベクトル量子化部 272 及び予測フィルタ 274 に供給する。

ベクトル量子化部 272 は、線形予測係数を要素とするコードベクトルとコードとを対応付けたコードブックを記憶しており、そのコードブックに基づいて、LPC 分析部 271 からの注目フレームの線形予測係数で構成される特徴ベクトルをベクトル量子化し、そのベクトル量子化の結果得られる A コードを、フィルタ係数復号器 273、並びにタップ生成部 278 及び 279 に供給する。

フィルタ係数復号器 273 は、ベクトル量子化部 272 が記憶しているのと同一のコードブックを記憶しており、そのコードブックに基づいて、ベクトル量子化部 272 からの A コードを、線形予測係数に復号し、音声合成フィルタ 277 に供給する。ここで、図 24 のフィルタ係数復号器 242 と、図 27 のフィルタ係数復号器 273 とは、同一構成のものとなっている。

予測フィルタ 274 は、そこに供給される注目フレームの音声信号と、LPC

分析部 271 からの線形予測係数を用いて、例えば、前述した式(1)にしたがった演算を行うことにより、注目フレームの残差信号を求め、ベクトル量子化部 275 に供給する。

即ち、式(1)における s_n と e_n の Z 変換を、S と E とそれぞれ表すと、式(1)は、次式のように表すことができる。

$$E = (1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots + \alpha_p z^{-p}) S \quad \dots \quad (16)$$

式(14)から、残差信号 e を求める予測フィルタ 274 は、FIR(Finite Impulse Response)型のデジタルフィルタで構成することができる。

即ち、図 28 は、予測フィルタ 274 の構成例を示している。

予測フィルタ 274 には、LPC 分析部 271 から、P 次の線形予測係数が供給されるようになっており、従って、予測フィルタ 274 は、P 個の遅延回路(D) 291, 乃至 291, P 個の乗算器 292, 乃至 292, 及び 1 つの加算器 293 から構成されている。

乗算器 292, 乃至 292, には、それぞれ、LPC 分析部 271 から供給される P 次の線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ がセットされる。

一方、注目フレームの音声信号 s は、遅延回路 291, と加算器 293 に供給される。遅延回路 291, は、そこへの入力信号を、残差信号の 1 サンプル分だけ遅延して、後段の遅延回路 291, に出力するとともに、演算器 292, に出力する。乗算器 292, は、遅延回路 291, の出力と、そこにセットされた線形予測係数 α_i とを乗算し、その乗算値を、加算器 293 に出力する。

加算器 293 は、乗算器 292, 乃至 292, の出力すべてと、音声信号 s とを加算し、その加算結果を、残差信号 e として出力する。

図 27 に戻り、ベクトル量子化部 275 は、残差信号のサンプル値を要素とするコードベクトルとコードとを対応付けたコードブックを記憶しており、そのコードブックに基づいて、予測フィルタ 274 からの注目フレームの残差信号のサンプル値で構成される残差ベクトルをベクトル量子化し、そのベクトル量子化の結果得られる残差コードを、残差コードブック記憶部 276, 並びにタップ生成部 278 及び 279 に供給する。

残差コードブック記憶部 276 は、ベクトル量子化部 275 が記憶しているの

と同一のコードブックを記憶しており、そのコードブックに基づいて、ベクトル量子化部 275 からの残差コードを、残差信号に復号し、音声合成フィルタ 277 に供給する。ここで、図 24 の残差コードブック記憶部 243 と、図 27 の残差コードブック記憶部 276 の記憶内容は同一になっている。

音声合成フィルタ 277 は、図 24 の音声合成フィルタ 244 と同様に構成される IIR フィルタで、フィルタ係数復号器 273 からの線形予測係数を IIR フィルタのタップ係数とともに、残差コードブック記憶部 276 からの残差信号を入力信号として、その入力信号のフィルタリングを行うことにより、合成音を生成し、タップ生成部 278 及び 279 に供給する。

タップ生成部 278 は、図 24 のタップ生成部 245 における場合と同様に、音声合成フィルタ 277 から供給される合成音、ベクトル量子化部 272 から供給される A コード、及びベクトル量子化部 275 から供給される残差コードから予測タップを構成し、正規方程式加算回路 281 に供給する。タップ生成部 279 は、図 24 のタップ生成部 246 における場合と同様に、音声合成フィルタ 277 から供給される合成音、ベクトル量子化部 272 供給される A コード、及びベクトル量子化部 275 から供給される残差コードから、クラスタップを構成し、クラス分類部 280 に供給する。

クラス分類部 280 は、図 24 のクラス分類部 247 における場合と同様に、そこに供給されるクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路 281 に供給する。

正規方程式加算回路 281 は、教師データとしての注目フレームの高音質の音声である学習用の音声と、タップ生成部 278 からの生徒データとしての予測タップを対象とした足し込みを行う。

即ち、正規方程式加算回路 281 は、クラス分類部 280 から供給されるクラスコードに対応するクラスごとに、予測タップ（生徒データ）を用い、前述した式（13）の行列 A における各コンポーネントとなっている、生徒データどうしの乗算 ($x_{ij}x_{ji}$) と、サメーション (Σ) に相当する演算を行う。

さらに、正規方程式加算回路 281 は、やはり、クラス分類部 280 から供給されるクラスコードに対応するクラスごとに、生徒データ及び教師データを用い、

式(13)のベクトルvにおける各コンポーネントとなっている、生徒データと教師データの乗算($x_i y_i$)と、サメーション(Σ)に相当する演算を行う。

正規方程式加算回路281は、以上の足し込みを、そこに供給される学習用の音声のフレームすべてを注目フレームとして行い、これにより、各クラスについて、式(13)に示した正規方程式をたてる。

タップ係数決定回路281は、正規方程式加算回路281においてクラスごとに生成された正規方程式を解くことにより、クラスごとに、タップ係数を求め、係数メモリ283の、各クラスに対応するアドレスに供給する。

なお、学習用の音声信号として用意した音声信号によっては、正規方程式加算回路281において、タップ係数を求めるのに必要な数の正規方程式が得られないクラスが生じる場合があり得るが、タップ係数決定回路281は、そのようなクラスについては、例えば、デフォルトのタップ係数を出力する。

係数メモリ283は、タップ係数決定回路281から供給されるクラスごとのタップ係数を、そのクラスに対応するアドレスに記憶する。

次に、図29のフローチャートを参照して、図27の学習装置の習処理について説明する。

学習装置には、学習用の音声信号が供給され、この学習用の音声信号は、LPC分析部271及び予測フィルタ274に供給されるとともに、教師データとして、正規方程式加算回路281に供給される。そして、ステップS211において、学習用の音声信号から、生徒データが生成される。

即ち、LPC分析部271は、学習用の音声信号のフレームを、順次、注目フレームとし、その注目フレームの音声信号をLPC分析することで、P次の線形予測係数を求め、ベクトル量子化部272に供給する。ベクトル量子化部272は、LPC分析部271からの注目フレームの線形予測係数で構成される特徴ベクトルをベクトル量子化し、そのベクトル量子化の結果得られるAコードを、生徒データとして、フィルタ係数復号器273、並びにタップ生成部278及び279に供給する。フィルタ係数復号器273は、ベクトル量子化部272からのAコードを、線形予測係数に復号し、その線形予測係数を、音声合成フィルタ277に供給する。

一方、LPC分析部271から注目フレームの線形予測係数を受信した予測フィルタ274は、その線形予測係数と、注目フレームの学習用の音声信号とを用いて、前述した式(1)にしたがった演算を行うことにより、注目フレームの残差信号を求め、ベクトル量子化部275に供給する。ベクトル量子化部275は、予測フィルタ274からの注目フレームの残差信号のサンプル値で構成される残差ベクトルをベクトル量子化し、そのベクトル量子化の結果得られる残差コードを、生徒データとして、残差コードブック記憶部276、並びにタップ生成部278及び279に供給する。残差コードブック記憶部276は、ベクトル量子化部275からの残差コードを、残差信号に復号し、音声合成フィルタ277に供給する。

以上のようにして、音声合成フィルタ277は、線形予測係数と残差信号を受信すると、その線形予測係数と残差信号を用いて音声合成を行い、その結果得られる合成音を、生徒データとして、タップ生成部278及び279に出力する。

そして、ステップS212に進み、タップ生成部278が、音声合成フィルタ277から供給される合成音、ベクトル量子化部272から供給されるAコード、及びベクトル量子化部275から供給される残差コードから、予測タップとクラスタップをそれぞれ生成する。予測タップは、正規方程式加算回路281に供給され、クラスタップは、クラス分類部280に供給される。

その後、ステップS213において、クラス分類部280が、タップ生成部279からのクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路281に供給する。

ステップS214に進み、正規方程式加算回路281は、クラス分類部280から供給されるクラスについて、そこに供給される教師データとしての注目フレームの高音質の音声のサンプル値、及びタップ生成部278からの生徒データとしての予測タップを対象とした、式(13)の行列Aとベクトルvの、上述したような足し込みを行い、ステップS215に進む。

ステップS215では、まだ、注目フレームとして処理すべきフレームの学習用の音声信号があるかどうかが判定される。ステップS215において、まだ、注目フレームとして処理すべきフレームの学習用の音声信号があると判定された

場合、ステップS211に戻り、次のフレームを新たに注目フレームとして、以下、同様の処理が繰り返される。

また、ステップS215において、注目フレームとして処理すべきフレームの学習用の音声信号がないと判定された場合、即ち、正規方程式加算回路281において、各クラスについて、正規方程式が得られた場合、ステップS216に進み、タップ係数決定回路281は、各クラスごとに生成された正規方程式を解くことにより、各クラスごとに、タップ係数を求め、係数メモリ283の、各クラスに対応するアドレスに供給して記憶させ、処理を終了する。

以上のようにして、係数メモリ283に記憶された各クラスごとのタップ係数が、図24の係数メモリ248に記憶されている。

従って、図3の係数メモリ248に記憶されたタップ係数は、線形予測演算を行うことにより得られる高音質の音声の予測値の予測誤差（ここでは、自乗誤差）が、統計的に最小になるように学習を行うことにより求められたものであるから、図24の予測部249が出力する音声は、音声合成フィルタ244で生成された合成音の歪みが低減（解消）された、高音質のものとなる。

なお、図24の音声合成装置において、上述したように、例えば、タップ生成部246に、線形予測係数や残差信号等の中からも、クラスタップを抽出させるようとする場合には、図27のタップ生成部278にも、図中点線で示すように、フィルタ係数復号器273が出力する線形予測係数や、残差コードブック記憶部276が出力する残差信号の中から、同様のクラスタップを抽出させるようとする必要がある。図24のタップ生成部245と、図27のタップ生成部278で生成される予測タップについても同様である。

上述の場合には、説明を簡単にするために、クラスタップを構成するビットの系列を、そのままクラスコードとするクラス分類を行うようにしたが、この場合、クラス数が膨大になることがある。そこで、クラス分類では、例えば、クラスタップをベクトル量子化等によって圧縮し、その圧縮の結果得られるビットの系列を、クラスコードとするようにすることが可能である。

次に、本発明を適用した伝送システムの一例を、図30を参照して説明する。ここで、システムとは、複数の装置が論理的に集合した物をいい、各構成の装置

が同一筐体中にあるか否かは問わない。

この伝送システムでは、携帯電話機401₁と401₂が、基地局402₁と402₂、それぞれとの間で、無線による送受信を行うとともに、基地局402₁と402₂、それそれが、交換局403との間で送受信を行うことにより、最終的には、携帯電話機401₁と401₂との間において、基地局402₁と402₂並びに交換局403を介して、音声の送受信を行うことができるようになっている。なお、基地局402₁と402₂は、同一の基地局であってもよいし、異なる基地局であってもよい。

ここで、以下、特に区別する必要がない限り、携帯電話機401₁と401₂を、携帯電話機401と記述する。

図30に示す携帯電話機401の具体的な構成を図31に示す。

アンテナ411は、基地局402₁と402₂からの電波を受信し、その受信信号を、変復調部412に供給するとともに、変復調部412からの信号を、電波で、基地局402₁又は402₂に送信する。変復調部412は、アンテナ411からの信号を復調し、その結果得られる、図1で説明したようなコードデータを、受信部414に供給する。また、変復調部412は、送信部413から供給される、図1で説明したようなコードデータを変調し、その結果得られる変調信号をアンテナ411に供給する。送信部413は、図1に示した送信部と同様に構成され、そこに入力されるユーザの音声を、コードデータに符号化して、変復調部412に供給する。受信部414は、変復調部412からのコードデータを受信し、そのコードデータから、図24の音声合成装置における場合と同様の高音質の音声を復号して出力する。

即ち、図31に示す携帯電話機401の受信部414の具体的な構成例を図32に示す。なお、図中、前述した図2における場合と対応する部分については、同一の符号を付してあり、以下では、その説明は、適宜省略する。

タップ生成部221と222には、音声合成フィルタ29が出力する、フレームごとの合成音と、チャネルデコーダ21が出力する、フレーム又はサブフレームごとのLコード、Gコード、Iコード、及びAコードとが供給されるようになっている。タップ生成部221と222は、そこに供給される合成音、Lコード、

Gコード、Iコード、及びAコードから、予測タップとするものとクラスタップとするものをそれぞれ抽出する。予測タップは、予測部225に供給され、クラスタップは、クラス分類部223に供給される。

クラス分類部223は、タップ生成部122から供給されるクラスタップに基づいて、クラス分類を行い、そのクラス分類結果としてのクラスコードを、係数メモリ224に供給する。

係数メモリ224は、後述する図33の学習装置において学習処理が行われることにより得られる、クラスごとのタップ係数を記憶しており、クラス分類部223が出力するクラスコードに対応するアドレスに記憶されているタップ係数を、予測部225に供給する。

予測部225は、図24の予測部249と同様に、タップ生成部221が出力する予測タップと、係数メモリ224が出力するタップ係数とを取得し、その予測タップとタップ係数とを用いて、前述した式(6)に示した線形予測演算を行う。これにより、予測部225は、注目フレームの高音質の音声の予測値求めて、D/A変換部30に供給する。

以上のように構成される受信部414では、基本的には、図26に示したフローチャートにしたがった処理と同様の処理が行われることで、高音質の合成音が、音声の復号結果として出力される。

即ち、チャネルデコーダ21は、そこに供給されるコードデータから、Lコード、Gコード、Iコード、Aコードを分離し、それを、適応コードブック記憶部22、ゲイン復号器23、励起コードブック記憶部24、フィルタ係数復号器25に供給する。さらに、Lコード、Gコード、Iコード、及びAコードは、タップ生成部221及び222にも供給される。

適応コードブック記憶部22、ゲイン復号器23、励起コードブック記憶部24、演算器26乃至28では、図1の適応コードブック記憶部9、ゲイン復号器10、励起コードブック記憶部11、演算器12乃至14における場合と同様の処理が行われ、これにより、Lコード、Gコード、及びIコードが、残差信号eに復号される。この残差信号は、音声合成フィルタ29に供給される。

さらに、フィルタ係数復号器25は、図1で説明したように、そこに供給され

るAコードを、線形予測係数に復号し、音声合成フィルタ29に供給する。音声合成フィルタ29は、演算器28からの残差信号と、フィルタ係数復号器25からの線形予測係数を用いて音声合成を行い、その結果得られる合成音を、タップ生成部221と222に供給する。

タップ生成部221は、音声合成フィルタ29が output する合成音のフレームを注目フレームとし、ステップS201において、その注目フレームの合成音と、Lコード、Gコード、Iコード、及びAコードとから、予測タップを生成し、予測部225に供給する。さらに、ステップS201では、タップ生成部222は、やはり、注目フレームの合成音と、Lコード、Gコード、Iコード、及びAコードとから、クラスタップを生成し、クラス分類部223に供給する。

そして、ステップS202に進み、クラス分類部223は、タップ生成部222から供給されるクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、係数メモリ224に供給して、ステップS203に進む。

ステップS203では、係数メモリ224は、クラス分類部223から供給されるクラスコードに対応するアドレスから、タップ係数を読み出し、予測部225に供給する。

ステップS204に進み、予測部225は、係数メモリ224が output するタップ係数を取得し、そのタップ係数と、タップ生成部221からの予測タップとを用いて、式(6)に示した積和演算を行い、注目フレームの高音質の音声の予測値を得る。

以上のようにして得られた高音質の音声は、予測部225から、D/A変換部30を介して、スピーカ31に供給され、これにより、スピーカ31からは、高音質の音声が output される。

ステップS204の処理後は、ステップS205に進み、まだ、注目フレームとして処理すべきフレームがあるかどうかが判定され、あると判定された場合、ステップS201に戻り、次に注目フレームとすべきフレームを、新たに注目フレームとして、以下、同様の処理を繰り返す。また、ステップS205において、注目フレームとして処理すべきフレームがないと判定された場合、処理を終了する。

次に、図32の係数メモリ224に記憶させるタップ係数の学習処理を行う学習装置の一例を図33を参照して説明する。

マイクロフォン501乃至コード決定部515は、図1のマイクロフォン1乃至コード決定部515とそれ同様に構成される。マイクロフォン501には、学習用の音声信号が入力されるようになっており、従って、マイク501乃至コード決定部515では、その学習用の音声信号に対して、図1における場合と同様の処理が施される。

そして、タップ生成部431と432には、自乗誤差最小判定部508において自乗誤差が最小になったと判定されたときの音声合成フィルタ506が出力する合成音が供給される。さらに、タップ生成部431と432には、コード決定部515が、自乗誤差最小判定部508から確定信号を受信したときに出力するLコード、Gコード、Iコード、及びAコードも供給される。また、正規方程式加算回路434には、A/D変換部202が出力する音声が、教師データとして供給される。

タップ生成部431は、音声合成フィルタ506が出力する合成音と、コード決定部515が出力するLコード、Gコード、Iコード、及びAコードとから、図32のタップ生成部221と同一の予測タップを構成し、生徒データとして、正規方程式加算回路234に供給する。

タップ生成部232も、音声合成フィルタ506が出力する合成音と、コード決定部515が出力するLコード、Gコード、Iコード、及びAコードとから、図32のタップ生成部222と同一のクラスタップを構成し、クラス分類部433に供給する。

クラス分類部433は、タップ生成部432からのクラスタップに基づいて、図32のクラス分類部223における場合と同様のクラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路434に供給する。

正規方程式加算回路434は、A/D変換部502からの音声を、教師データとして受信するとともに、タップ生成部131からの予測タップを、生徒データとして受信し、その教師データ及び生徒データを対象として、クラス分類部433からのクラスコードごとに、図27の正規方程式加算回路281における場合

と同様の足し込みを行うことにより、各クラスについて、式(13)に示した正規方程式をたてる。

タップ係数決定回路435は、正規方程式加算回路434においてクラスごとに生成された正規方程式を解くことにより、クラスごとに、タップ係数を求め、係数メモリ436の、各クラスに対応するアドレスに供給する。

なお、学習用の音声信号として用意する音声信号によっては、正規方程式加算回路434において、タップ係数を求めるのに必要な数の正規方程式が得られないクラスが生じる場合があり得るが、タップ係数決定回路435は、そのようなクラスについては、例えば、デフォルトのタップ係数を出力する。

係数メモリ436は、タップ係数決定回路435から供給されるクラスごとの線形予測係数と残差信号についてのタップ係数を記憶する。

以上のように構成される学習装置では、基本的には、図29に示したフローチャートにしたがった処理と同様の処理が行われることで、高音質の合成音を得るためにタップ係数が求められる。

即ち、学習装置には、学習用の音声信号が供給され、ステップS211では、その学習用の音声信号から、教師データと生徒データが生成される。

即ち、学習用の音声信号は、マイクロフォン501に入力され、マイクロフォン501乃至コード決定部515は、図1のマイクロフォン1乃至コード決定部15における場合とそれぞれ同様の処理を行う。

その結果、A/D変換部502で得られるデジタル信号の音声は、教師データとして、正規方程式加算回路434に供給される。また、自乗誤差最小判定部508において自乗誤差が最小になったと判定されたときに、音声合成フィルタ506が出力する合成音は、生徒データとして、タップ生成部431と432に供給される。さらに、自乗誤差最小判定部208において自乗誤差が最小になったと判定されたときに、コード決定部515が出力するLコード、Gコード、Iコード、及びAコードも、生徒データとして、タップ生成部431と432に供給される。

その後、ステップS212に進み、タップ生成部431は、音声合成フィルタ506から生徒データとして供給される合成音のフレームを注目フレームとして、

その注目フレームの合成音と、Lコード、Gコード、Iコード、及びAコードとから、予測タップを生成し、正規方程式加算回路434に供給する。さらに、ステップS212では、タップ生成部432が、やはり、注目フレームの合成音と、Lコード、Gコード、Iコード、及びAコードとから、クラスタップを生成し、クラス分類部433に供給する。

ステップS212の処理後は、ステップS213に進み、クラス分類部433が、タップ生成部432からのクラスタップに基づいて、クラス分類を行い、その結果得られるクラスコードを、正規方程式加算回路434に供給する。

ステップS214に進み、正規方程式加算回路434は、A/D変換器502からの教師データとしての注目フレームの高音質の音声である学習用の音声、及びタップ生成部432からの生徒データとしての予測タップを対象として、式(13)の行列Aとベクトルvの、上述したような足し込みを、クラス分類部433からのクラスコードごとに行い、ステップS215に進む。

ステップS215では、まだ、注目フレームとして処理すべきフレームがあるかどうかが判定される。ステップS215において、まだ、注目フレームとして処理すべきフレームがあると判定された場合、ステップS211に戻り、次のフレームを新たに注目フレームとして、以下、同様の処理が繰り返される。

また、ステップS215において、注目フレームとして処理すべきフレームがないと判定された場合、即ち、正規方程式加算回路434において、各クラスについて、正規方程式が得られた場合、ステップS216に進み、タップ係数決定回路435は、各クラスごとに生成された正規方程式を解くことにより、各クラスごとに、タップ係数を求め、係数メモリ436の、各クラスに対応するアドレスに供給して記憶させ、処理を終了する。

以上のようにして、係数メモリ436に記憶された各クラスごとのタップ係数が、図32の係数メモリ224に記憶されている。

従って、図32の係数メモリ224に記憶されたタップ係数は、線形予測演算を行うことにより得られる高音質の音声予測値の予測誤差(自乗誤差)が、統計的に最小になるように学習を行うことにより求められたものであるから、図32の予測部225が出力する音声は高音質のものとなる。

図32及び図33に示す例では、クラスタップを、音声合成フィルタ506が
出力する合成音と、Lコード、Gコード、Iコード、及びAコードとから生成
するようにしたが、クラスタップは、Lコード、Gコード、Iコード、又はAコ
ードのうちの1以上と、音声合成フィルタ506が出力する合成音とから生成す
ることが可能である。また、クラスタップは、図32において点線で示すように、
Aコードから得られる線形予測係数 α や、Gコードから得られるゲイン β 、 γ 、
その他の、Lコード、Gコード、Iコード、又はAコードから得られる情報、例
えば、残差信号eや、残差信号eを得るためのl, n、さらには、 $1/\beta$, n/γ などをも用いて構成することが可能である。さらに、クラスタップは、音声合
成フィルタ506が出力する合成音と、Lコード、Gコード、Iコード、又はA
コードから得られる上述したような情報とから生成することも可能である。また、
CELP方式では、コードデータに、リスト補間ビットやフレームエネルギーが含
められる場合があるが、この場合、クラスタップは、ソフト補間ビットやフレー
ムエネルギーを用いて構成することが可能である。予測タップについても、同様で
ある。

ここで、図34に、図33の学習装置において、教師データとして用いられる
音声データsと、生徒データとして用いられる合成音のデータss、残差信号e、
残差信号を求めるのに用いられるn及びlを示す。

次に、上述した一連の処理は、ハードウェアにより行うこともできるし、ソフ
トウェアにより行うこともできる。一連の処理をソフトウェアによって行う場合
には、そのソフトウェアを構成するプログラムが、汎用のコンピュータ等にイン
ストールされる。

上述した一連の処理は、ハードウェアにより行うこともできるし、ソフトウェ
アにより行うこともできる。一連の処理をソフトウェアによって行う場合には、
そのソフトウェアを構成するプログラムが、汎用のコンピュータ等にインストー
ルされる。

上述した一連の処理を実行するプログラムがインストールされるコンピュータ
は、前述した図13に示すように構成され、図13に示したコンピュータと同様
の動作が実行されるので、その詳細な説明は省略する。

本発明において、コンピュータに各種の処理を行わせるためのプログラムを記述する処理ステップは、必ずしもフローチャートとして記載された順序に沿って時系列に処理する必要はなく、並列的あるいは個別に実行される処理（例えば、並列処理あるいはオブジェクトによる処理）も含むものである。

また、プログラムは、1のコンピュータにより処理されるものであっても良いし、複数のコンピュータによって分散処理されるものであっても良い。さらに、プログラムは、遠方のコンピュータに転送されて実行されるものであっても良い。

なお、本例においても、学習用の音声信号として、どのようなものを用いるかについては、特に言及しなかったが、学習用の音声信号としては、人が発話した音声の他、例えば、曲（音楽）等を採用することが可能である。上述したような学習処理によれば、学習用の音声信号として、人の発話を用いた場合には、そのような人の発話の音声の音質を向上させるようなタップ係数が得られ、曲を用いた場合には、曲の音質を向上させるようなタップ係数が得されることになる。

さらに、本発明は、例えば、V S E L P (Vector Sum Excited Liner Prediction), P S I - C E L P (Pitch Synchronous Innovation CELP), C S - A C E L P (Conjugate Structure Algebraic CELP) 等の C E L P 方式による符号化の結果得られるコードから合成音を生成する場合に、広く適用可能である。

また、本発明は、C E L P 方式による符号化の結果得られるコードから合成音を生成する場合に限らず、あるコードから、残差信号と線形予測係数を得て、合成音を生成する場合に、広く適用可能である。

さらに、上述の説明では、タップ係数を用いた線形1次予測演算によって、残差信号や線形予測係数の予測値を求めるようにしたが、この予測値は、その他、2次以上の高次の予測演算によって求めることも可能である。

また、上述の説明では、クラスタップをベクトル量子化すること等によってクラス分類を行うようにしたが、クラス分類は、その他、例えば、A D R C 処理を利用して行うこと等も可能である。

A D R C を用いるクラス分類では、クラスタップを構成する要素、すなわち、合成音のサンプル値や、Lコード、Gコード、Iコード、Aコード等が、A D R C 処理され、その結果得られるADRCコードにしたがって、クラスが決定される。

ここで、KビットADRCにおいては、例えば、クラスタップを構成する要素の最大値MAXと最小値MINが検出され、 $DR=MAX-MIN$ を、集合の局所的なダイナミックレンジとし、このダイナミックレンジDRに基づいて、クラスタップを構成する要素がKビットに再量子化される。即ち、クラスタップを構成する各要素から、最小値MINが減算され、その減算値が $DR/2^K$ で量子化される。そして、以上のようにして得られる、クラスタップを構成する各要素のKビットの値を、所定の順番で並べたビット列が、ADRCコードとして出力される。

産業上の利用可能性

上述したように、本発明は、予測値を求めようとしている高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップが、合成音と、コード又はコードから得られる情報とから抽出されるとともに、注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップが、合成音と、コード又はコードから得られる情報とから抽出され、クラスタップに基づいて、注目音声のクラスを求めるクラス分類が行われ、予測タップと、注目音声のクラスに対応するタップ係数を用いて、注目音声の予測値を求ることにより、高音質の合成音を生成することが可能となる。

請求の範囲

1. 所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音から、その音質を向上させた高音質の音声の予測値を予測するための予測タップを抽出し、その予測タップと所定のタップ係数を用いて、所定の予測演算を行うことにより前記高音質の音声の予測値を求める音声処理装置であって、

前記予測値を求めようとしている前記高音質の音声を注目音声として、その注目音声を予測するのに用いる前記予測タップを、前記合成音から抽出する予測タップ抽出手段と、

前記注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記コードから抽出するクラスタップ抽出手段と、

前記クラスタップに基づいて前記注目音声のクラスを求めるクラス分類を行うクラス分類手段と、

学習を行うことにより求められた前記クラスごとの前記タップ係数の中から前記注目音声のクラスに対応する前記タップ係数を取得する取得手段と、

前記予測タップと、前記注目音声のクラスに対応する前記タップ係数を用いて前記注目音声の予測値を求める予測手段とを備えることを特徴とするデータ処理装置。

2. 前記予測手段は、前記予測タップ及びタップ係数を用いて線形1次予測演算を行うことにより、前記注目音声の予測値を求める特徴とする請求の範囲第1項記載のデータ処理装置。

3. 前記取得手段は、クラスごとの前記タップ係数を記憶している記憶手段から、前記注目音声に対応するクラスの前記タップ係数を取得することを特徴とする請求の範囲第1項記載のデータ処理装置。

4. 前記クラスタップ抽出手段は、前記クラスタップを、前記コードと、そのコードを復号することにより得られる前記線形予測係数又は残差信号の中から抽出することを特徴とする請求の範囲第1項記載のデータ処理装置。

5. 前記タップ係数は、前記予測タップ及びタップ係数を用いて所定の予測演算

を行うことにより得られる前記高音質の音声の予測値の予測誤差が、統計的に最小になるように、学習を行うことにより得られたものであることを特徴とする請求の範囲第1項記載のデータ処理装置。

6. 前記音声合成フィルタをさらに備えることを特徴とする請求の範囲第1項記載のデータ処理装置。

7. 前記コードは、音声を、CELP(Code Excited Liner Prediction coding)方式によって符号化することにより得られたものであることを特徴とする請求の範囲第1項記載のデータ処理装置。

8. 所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音から、その音質を向上させた高音質の音声の予測値を予測するための予測タップを抽出し、その予測タップと所定のタップ係数を用いて、所定の予測演算を行うことにより、前記高音質の音声の予測値を求める音声処理方法であって、

前記予測値を求めようとしている前記高音質の音声を注目音声として、その注目音声を予測するのに用いる前記予測タップを、前記合成音から抽出する予測タップ抽出ステップと、

前記注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記コードから抽出するクラスタップ抽出ステップと、

前記クラスタップに基づいて、前記注目音声のクラスを求めるクラス分類を行うクラス分類ステップと、

学習を行うことにより求められた、前記クラスごとの前記タップ係数の中から前記注目音声のクラスに対応する前記タップ係数を取得する取得ステップと、

前記予測タップと、前記注目音声のクラスに対応する前記タップ係数を用いて、前記注目音声の予測値を求める予測ステップとを備えることを特徴とするデータ処理方法。

9. 所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音から、その音質を向上させた高音質の音声の予測値を予測するための予測タップを抽出し、その予測タップと所定のタップ係数を用いて、所定の予測演算を行うことにより、前記高音質の音声の予測値を

求める音声処理を、コンピュータに行わせるプログラムが記録されている記録媒体であって、

前記予測値を求めようとしている前記高音質の音声を注目音声として、その注目音声を予測するのに用いる前記予測タップを、前記合成音から抽出する予測タップ抽出ステップと、

前記注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記コードから抽出するクラスタップ抽出ステップと、

前記クラスタップに基づいて、前記注目音声のクラスを求めるクラス分類を行うクラス分類ステップと、

学習を行うことにより求められた、前記クラスごとの前記タップ係数の中から、前記注目音声のクラスに対応する前記タップ係数を取得する取得ステップと、

前記予測タップと、前記注目音声のクラスに対応する前記タップ係数を用いて、前記注目音声の予測値を求める予測ステップとを備えるプログラムが記録されていることを特徴とする記録媒体。

10. 所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音から、その音質を向上させた高音質の音声の予測値を、所定の予測演算によって求めるのに用いる所定のタップ係数を学習する学習装置であって、

前記予測値を求めようとしている前記高音質の音声を注目音声として、その注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記コードから抽出するクラスタップ抽出手段と、

前記クラスタップに基づいて、前記注目音声のクラスを求めるクラス分類を行うクラス分類手段と、

前記タップ係数及び合成音を用いて予測演算を行うことにより得られる前記高音質の音声の予測値の予測誤差が統計的に最小になるように、学習を行い、前記クラスごとのタップ係数を求める学習手段とを備えることを特徴とする学習装置。

11. 前記学習手段は、前記タップ係数及び合成音を用いて線形1次予測演算を行うことにより得られる前記高音質の音声の予測値の予測誤差が、統計的に最小になるように学習を行うことを特徴とする請求の範囲第10項記載の学習装置。

12. 前記クラスタップ抽出手段は、前記クラスタップを、前記コードと、そのコードを復号することにより得られる前記線形予測係数又は残差信号の中から抽出することを特徴とする請求の範囲第10項記載の学習装置。

13. 前記コードは、音声を、C E L P (Code Excited Liner Prediction coding) 方式によって符号化することにより得られたものであることを特徴とする請求の範囲第10項記載の学習装置。

14. 所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音から、その音質を向上させた高音質の音声の予測値を、所定の予測演算によって求めるのに用いる所定のタップ係数を学習する学習方法であって、

前記予測値を求めようとしている前記高音質の音声を注目音声として、その注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記コードから抽出するクラスタップ抽出ステップと、

前記グラスタップに基づいて、前記注目音声のクラスを求めるクラス分類を行うクラス分類ステップと、

前記タップ係数及び合成音を用いて予測演算を行うことにより得られる前記高音質の音声の予測値の予測誤差が統計的に最小になるように、学習を行い、前記クラスごとのタップ係数を求める学習ステップとを備えることを特徴とする学習方法。

15. 所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音から、その音質を向上させた高音質の音声の予測値を、所定の予測演算によって求めるのに用いる所定のタップ係数を学習する学習処理を、コンピュータに行わせるプログラムが記録されている記録媒体であって、

前記予測値を求めようとしている前記高音質の音声を注目音声として、その注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記コードから抽出するクラスタップ抽出ステップと、

前記グラスタップに基づいて、前記注目音声のクラスを求めるクラス分類を行うクラス分類ステップと、

前記タップ係数及び合成音を用いて予測演算を行うことにより得られる前記高音質の音声の予測値の予測誤差が統計的に最小になるように、学習を行い、前記クラスごとのタップ係数を求める学習ステップとを備えるプログラムが記録されていることを特徴とする記録媒体。

16. 線形予測係数と所定の入力信号に基づいて音声合成を行う音声合成フィルタに与えるフィルタデータを、所定のコードから生成するデータ処理装置であつて、

前記コードを復号し、復号フィルタデータを出力するコード復号手段と、
学習を行うことにより求められた所定のタップ係数を取得する取得手段と、
前記タップ係数及び復号フィルタデータを用いて、所定の予測演算を行うこと
により、前記フィルタデータの予測値を求め、前記音声合成フィルタに供給する
予測手段とを備えることを特徴とするデータ処理装置。

17. 前記予測手段は、前記タップ係数及び復号フィルタデータを用いて線形1次予測演算を行うことにより、前記フィルタデータの予測値を求めることを特徴とする請求の範囲第16項記載のデータ処理装置。

18. 前記取得手段は、前記タップ係数を記憶している記憶手段から、前記タップ係数を取得することを特徴とする請求の範囲第16項記載のデータ処理装置。

19. 前記予測値を求めようとしている前記フィルタデータを注目フィルタデータとして、その注目フィルタデータを予測するのに前記タップ係数とともに用いる予測タップを、前記復号フィルタデータから抽出する予測タップ抽出手段をさらに備え、前記予測手段は、前記予測タップ及びタップ係数を用いて予測演算を行うことを特徴とする請求の範囲第16項記載のデータ処理装置。

20. 上記装置は、さらに、前記注目フィルタデータを、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記復号フィルタデータから抽出するクラスタップ抽出手段と、前記クラスタップに基づいて、前記注目フィルタデータのクラスを求めるクラス分類を行うクラス分類手段とを備え、

前記予測手段は、前記予測タップと、前記注目フィルタデータのクラスに対応する前記タップ係数とを用いて予測演算を行うことを特徴とする請求の範囲第19項記載のデータ処理装置。

2 1. 上記装置は、さらに、前記注目フィルタデータを、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記コードから抽出するクラスタップ抽出手段と、前記クラスタップに基づいて、前記注目フィルタデータのクラスを求めるクラス分類を行うクラス分類手段とを備え、

前記予測手段は、前記予測タップと、前記注目フィルタデータのクラスに対応する前記タップ係数とを用いて予測演算を行うことを特徴とする請求の範囲第19項記載のデータ処理装置。

2 2. 前記クラスタップ抽出手段は、前記クラスタップを、前記コードと、前記復号フィルタデータの両方から抽出することを特徴とする請求の範囲第21項記載記載のデータ処理装置。

2 3. 前記タップ係数は、前記タップ係数及び復号フィルタデータを用いて所定の予測演算を行うことにより得られる前記フィルタデータの予測値の予測誤差が、統計的に最小になるように、学習を行うことにより得られたものであることを特徴とする請求の範囲第16項記載のデータ処理装置。

2 4. 前記フィルタデータは、前記入力信号と線形予測係数のうちの少なくとも一方又は両方であることを特徴とする請求の範囲第16項記載のデータ処理装置。

2 5. 前記音声合成フィルタをさらに備えることを特徴とする請求の範囲第16項記載のデータ処理装置。

2 6. 前記コードは、音声を、C E L P (Code Excited Liner Prediction coding) 方式によって符号化することにより得られたものであることを特徴とする請求の範囲第16項記載のデータ処理装置。

2 7. 線形予測係数と所定の入力信号に基づいて音声合成を行う音声合成フィルタに与えるフィルタデータを、所定のコードから生成するデータ処理方法であつて、

前記コードを復号し、復号フィルタデータを出力するコード復号ステップと、学習を行うことにより求められた所定のタップ係数を取得する取得ステップと、前記タップ係数及び復号フィルタデータを用いて、所定の予測演算を行うことにより、前記フィルタデータの予測値を求め、前記音声合成フィルタに供給する予測ステップとを備えることを特徴とするデータ処理方法。

28. 線形予測係数と所定の入力信号に基づいて音声合成を行う音声合成フィルタに与えるフィルタデータを、所定のコードから生成するデータ処理を、コンピュータに行わせるプログラムが記録されている記録媒体であって、

前記コードを復号し、復号フィルタデータを出力するコード復号ステップと、学習を行うことにより求められた所定のタップ係数を取得する取得ステップと、前記タップ係数及び復号フィルタデータを用いて、所定の予測演算を行うことにより、前記フィルタデータの予測値を求め、前記音声合成フィルタに供給する予測ステップとを備えるプログラムが記録されていることを特徴とする記録媒体。

29. 線形予測係数と所定の入力信号に基づいて音声合成を行う音声合成フィルタに与えるフィルタデータに対応するコードから、前記フィルタデータの予測値を、予測演算によって求めるのに用いる所定のタップ係数を学習する学習装置であって、

フィルタデータに対応するコードを復号し、復号フィルタデータを出力するコード復号手段と、

前記タップ係数及び復号フィルタデータを用いて予測演算を行うことにより得られる前記フィルタデータの予測値の予測誤差が、統計的に最小になるように学習を行い、前記タップ係数を求める学習手段とを備えることを特徴とする学習装置。

30. 前記学習手段は、前記タップ係数及び復号フィルタデータを用いて線形1次予測演算を行うことにより得られる前記フィルタデータの予測値の予測誤差が、統計的に最小になるように学習を行うことを特徴とする請求の範囲第29項記載の学習装置。

31. 上記装置は、さらに、前記予測値を求めようとしている前記フィルタデータを注目フィルタデータとして、その注目フィルタデータを予測するのに前記タップ係数とともに用いる予測タップを、前記復号フィルタデータから抽出する予測タップ抽出手段を備え、

前記学習手段は、前記予測タップ及びタップ係数を用いて予測演算を行うことにより得られる前記フィルタデータの予測値の予測誤差が、統計的に最小になる

ように学習を行うことを特徴とする請求の範囲第29項記載の学習装置。

32. 上記装置は、さらに、前記注目フィルタデータを、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記復号フィルタデータから抽出するクラスタップ抽出手段と、前記クラスタップに基づいて、前記注目フィルタデータのクラスを求めるクラス分類を行うクラス分類手段とを備え、

前記学習手段は、前記予測タップと、前記注目フィルタデータのクラスに対応する前記タップ係数とを用いて予測演算を行うことにより得られる前記フィルタデータの予測値の予測誤差が、統計的に最小になるように学習を行うことを特徴とする請求の範囲第31項記載の学習装置。

33. 上記装置は、さらに、前記注目フィルタデータを、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記コードから抽出するクラスタップ抽出手段と、前記クラスタップに基づいて、前記注目フィルタデータのクラスを求めるクラス分類を行うクラス分類手段とを備え、

前記学習手段は、前記予測タップと、前記注目フィルタデータのクラスに対応する前記タップ係数とを用いて予測演算を行うことにより得られる前記フィルタデータの予測値の予測誤差が、統計的に最小になるように学習を行うことを特徴とする請求の範囲第31項記載の学習装置。

34. 前記クラスタップ抽出手段は、前記クラスタップを、前記コードと、前記復号フィルタデータの両方から抽出することを特徴とする請求の範囲第33項記載の学習装置。

35. 前記フィルタデータは、前記入力信号と線形予測係数のうちの少なくとも一方又は両方であることを特徴とする請求の範囲第29項記載の学習装置。

36. 前記コードは、音声を、CELP(Code Excited Liner Prediction coding)方式によって符号化することにより得られたものであることを特徴とする請求の範囲第29項記載の学習装置。

37. 線形予測係数と所定の入力信号に基づいて音声合成を行う音声合成フィルタに与えるフィルタデータに対応するコードから、前記フィルタデータの予測値を、予測演算によって求めるのに用いる所定のタップ係数を学習する学習方法であって、

フィルタデータに対応するコードを復号し、復号フィルタデータを出力するコード復号ステップと、

前記タップ係数及び復号フィルタデータを用いて予測演算を行うことにより得られる前記フィルタデータの予測値の予測誤差が、統計的に最小になるように学習を行い、前記タップ係数を求める学習ステップとを備えることを特徴とする学習方法。

38. 線形予測係数と所定の入力信号に基づいて音声合成を行う音声合成フィルタに与えるフィルタデータに対応するコードから、前記フィルタデータの予測値を、予測演算によって求めるのに用いる所定のタップ係数を学習する学習処理を、コンピュータに行わせるプログラムが記録されている記録媒体であって、

フィルタデータに対応するコードを復号し、復号フィルタデータを出力するコード復号ステップと、

前記タップ係数及び復号フィルタデータを用いて予測演算を行うことにより得られる前記フィルタデータの予測値の予測誤差が、統計的に最小になるように学習を行い、前記タップ係数を求める学習ステップとを備えるプログラムが記録されていることを特徴とする記録媒体。

39. 所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音から、その音質を向上させた高音質の音声の予測値を求める音声処理装置であって、

前記予測値を求めようとしている前記高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップを、前記合成音と、前記コード又は前記コードから得られる情報とから抽出する予測タップ抽出手段と、

前記注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記合成音と、前記コード又は前記コードから得られる情報とから抽出するクラスタップ抽出手段と、

前記クラスタップに基づいて前記注目音声のクラスを求めるクラス分類を行うクラス分類手段と、

学習を行うことにより求められた前記クラスごとの前記タップ係数の中から前記注目音声のクラスに対応する前記タップ係数を取得する取得手段と、

前記予測タップと前記注目音声のクラスに対応する前記タップ係数を用いて前記注目音声の予測値を求める予測手段とを備えることを特徴とするデータ処理装置。

4 0 . 前記予測手段は、前記予測タップ及びタップ係数を用いて線形1次予測演算を行うことにより、前記注目音声の予測値を求める特徴とする請求の範囲第39項記載のデータ処理装置。

4 1 . 前記取得手段は、クラスごとの前記タップ係数を記憶している記憶手段から、前記注目音声に対応するクラスの前記タップ係数を取得することを特徴とする請求の範囲第39項記載のデータ処理装置。

4 2 . 前記予測タップ抽出手段又はクラスタップ抽出手段は、前記予測タップ又はクラスタップを、前記合成音、前記コード、及びコードから得られる情報から抽出することを特徴とする請求の範囲第39項記載のデータ処理装置。

4 3 . 前記タップ係数は、前記予測タップ及びタップ係数を用いて所定の予測演算を行うことにより得られる前記高音質の音声の予測値の予測誤差が、統計的に最小になるように、学習を行うことにより得られたものであることを特徴とする請求の範囲第39項記載のデータ処理装置。

4 4 . 上記装置は、さらに、音声合成フィルタを備えることを特徴とする請求の範囲第39項記載のデータ処理装置。

4 5 . 前記コードは、音声を、CELP(Code Excited Liner Prediction coding)方式によって符号化することにより得られたものであることを特徴とする請求の範囲第39項記載のデータ処理装置。

4 6 . 所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音から、その音質を向上させた高音質の音声の予測値を求める音声処理方法であって、

前記予測値を求めようとしている前記高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップを、前記合成音と、前記コード又は前記コードから得られる情報とから抽出する予測タップ抽出ステップと、

前記注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記コード合成音と、前記コード又は前記コードから得られ

る情報とから抽出するクラスタップ抽出ステップと、

前記クラスタップに基づいて、前記注目音声のクラスを求めるクラス分類を行うクラス分類ステップと、

学習を行うことにより求められた、前記クラスごとの前記タップ係数の中から、前記注目音声のクラスに対応する前記タップ係数を取得する取得ステップと、

前記予測タップと、前記注目音声のクラスに対応する前記タップ係数を用いて、前記注目音声の予測値を求める予測ステップとを備えることを特徴とするデータ処理方法。

47. 所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音から、その音質を向上させた高音質の音声の予測値を求める音声処理を、コンピュータに行わせるプログラムが記録されている記録媒体であって、

前記予測値を求めようとしている前記高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップを、前記合成音と、前記コード又は前記コードから得られる情報とから抽出する予測タップ抽出ステップと、

前記注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記合成音と、前記コード又は前記コードから得られる情報とから抽出するクラスタップ抽出ステップと、

前記クラスタップに基づいて、前記注目音声のクラスを求めるクラス分類を行うクラス分類ステップと、

学習を行うことにより求められた、前記クラスごとの前記タップ係数の中から前記注目音声のクラスに対応する前記タップ係数を取得する取得ステップと、前記予測タップと前記注目音声のクラスに対応する前記タップ係数を用いて前記注目音声の予測値を求める予測ステップとを備えるプログラムが記録されていることを特徴とする記録媒体。

48. 所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音からその音質を向上させた高音質の音声の予測値を、所定の予測演算によって求めるのに用いる所定のタップ係数を学習する学習装置であって、

前記予測値を求めようとしている前記高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップを、前記合成音と、前記コード又は前記コードから得られる情報とから抽出する予測タップ抽出手段と、

前記注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記合成音と、前記コード又は前記コードから得られる情報とから抽出するクラスタップ抽出手段と、

前記クラスタップに基づいて、前記注目音声のクラスを求めるクラス分類を行うクラス分類手段と、

前記タップ係数及び予測タップを用いて予測演算を行うことにより得られる前記高音質の音声の予測値の予測誤差が統計的に最小になるように学習を行い、前記クラスごとのタップ係数を求める学習手段とを備えることを特徴とする学習装置。

4 9 . 前記学習手段は、前記タップ係数及び予測タップを用いて線形 1 次予測演算を行うことにより得られる前記高音質の音声の予測値の予測誤差が、統計的に最小になるように学習を行うことを特徴とする請求の範囲第 4 8 項記載の学習装置。

5 0 . 前記予測タップ抽出手段又はクラスタップ抽出手段は、前記予測タップ又はクラスタップを、前記合成音と、前記コード、及び前記コードから得られる情報から抽出することを特徴とする請求の範囲第 4 8 項記載の学習装置。

5 1 . 前記コードは、音声を、C E L P (Code Excited Liner Prediction coding) 方式によって符号化することにより得られたものであることを特徴とする請求の範囲第 4 8 項記載の学習装置。

5 2 . 所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音から、その音質を向上させた高音質の音声の予測値を、所定の予測演算によって求めるのに用いる所定のタップ係数を学習する学習方法であって、

前記予測値を求めようとしている前記高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップを、前記合成音と、前記コード又は前記コードから得られる情報とから抽出する予測タップ抽出手段と、

前記注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記合成音と、前記コード又は前記コードから得られる情報とから抽出するクラスタップ抽出ステップと、

前記クラスタップに基づいて、前記注目音声のクラスを求めるクラス分類を行うクラス分類ステップと、

前記タップ係数及び予測タップを用いて予測演算を行うことにより得られる前記高音質の音声の予測値の予測誤差が統計的に最小になるように、学習を行い、前記クラスごとのタップ係数を求める学習ステップとを備えることを特徴とする学習方法。

53. 所定のコードから生成される線形予測係数と残差信号を、音声合成フィルタに与えることによって得られる合成音から、その音質を向上させた高音質の音声の予測値を、所定の予測演算によって求めるのに用いる所定のタップ係数を学習する学習処理を、コンピュータに行わせるプログラムが記録されている記録媒体であって、

前記予測値を求めようとしている前記高音質の音声を注目音声として、その注目音声を予測するのに用いる予測タップを、前記合成音と、前記コード又は前記コードから得られる情報とから抽出する予測タップ抽出手段と、

前記注目音声を、幾つかのクラスのうちのいずれかにクラス分類するのに用いるクラスタップを、前記合成音と、前記コード又は前記コードから得られる情報とから抽出するクラスタップ抽出ステップと、

前記クラスタップに基づいて、前記注目音声のクラスを求めるクラス分類を行うクラス分類ステップと、

前記タップ係数及び予測タップを用いて予測演算を行うことにより得られる前記高音質の音声の予測値の予測誤差が統計的に最小になるように、学習を行い、前記クラスごとのタップ係数を求める学習ステップとを備えるプログラムが記録されていることを特徴とする記録媒体。

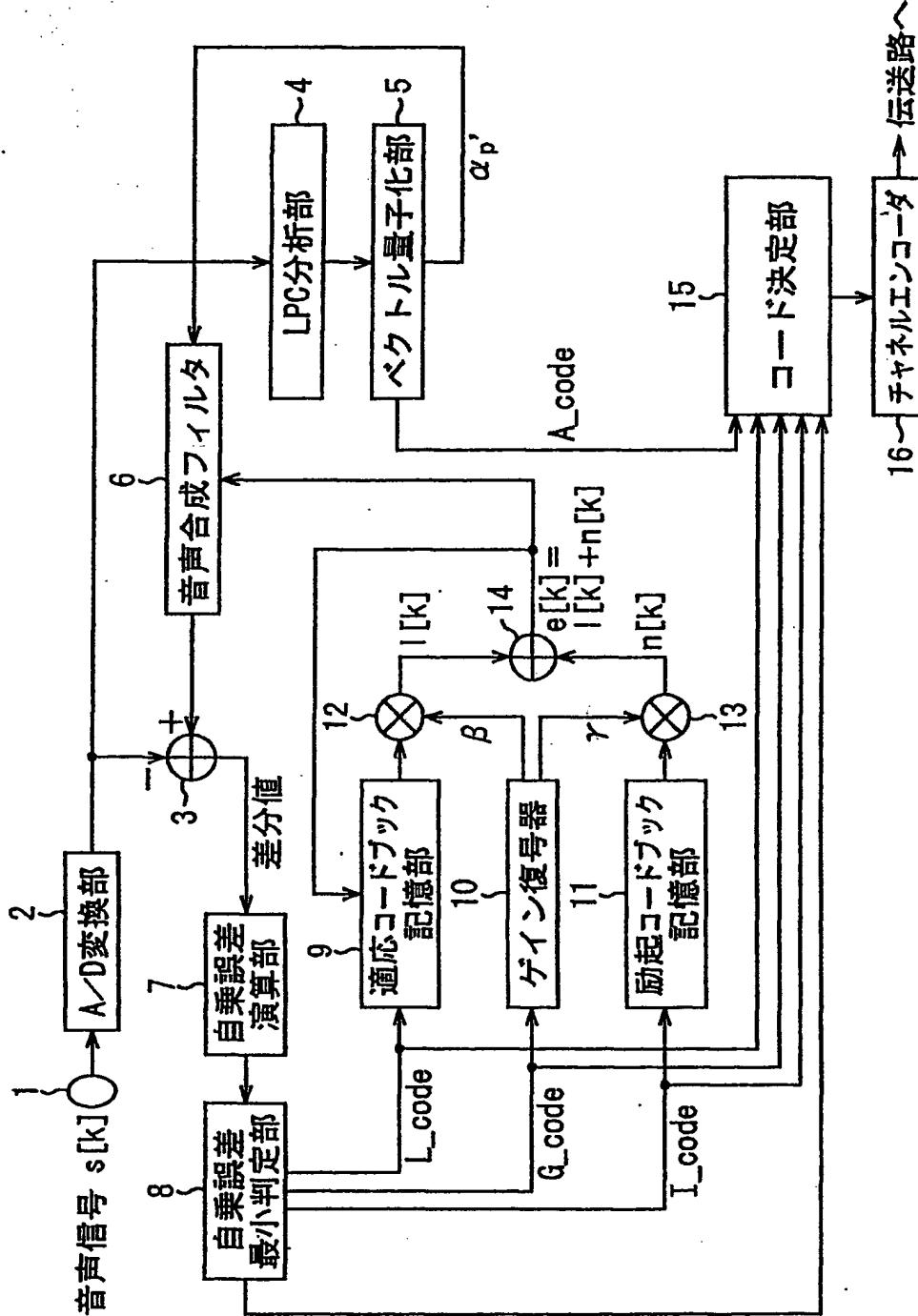


FIG.1

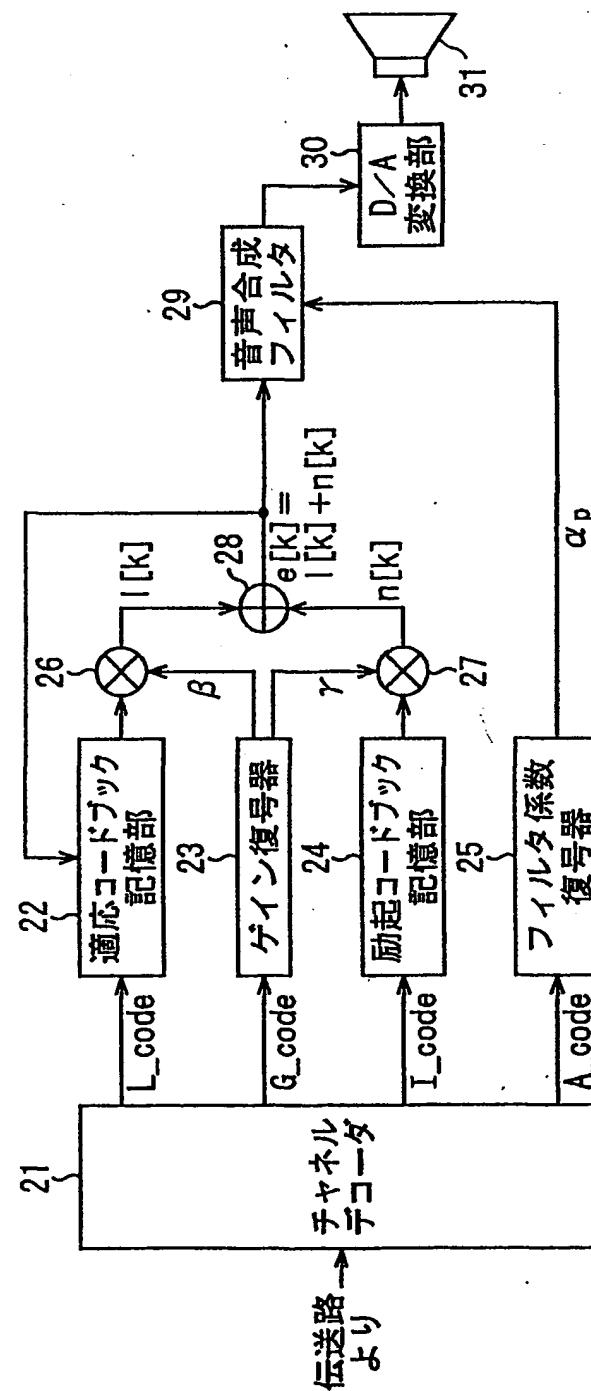


FIG.2

3/34

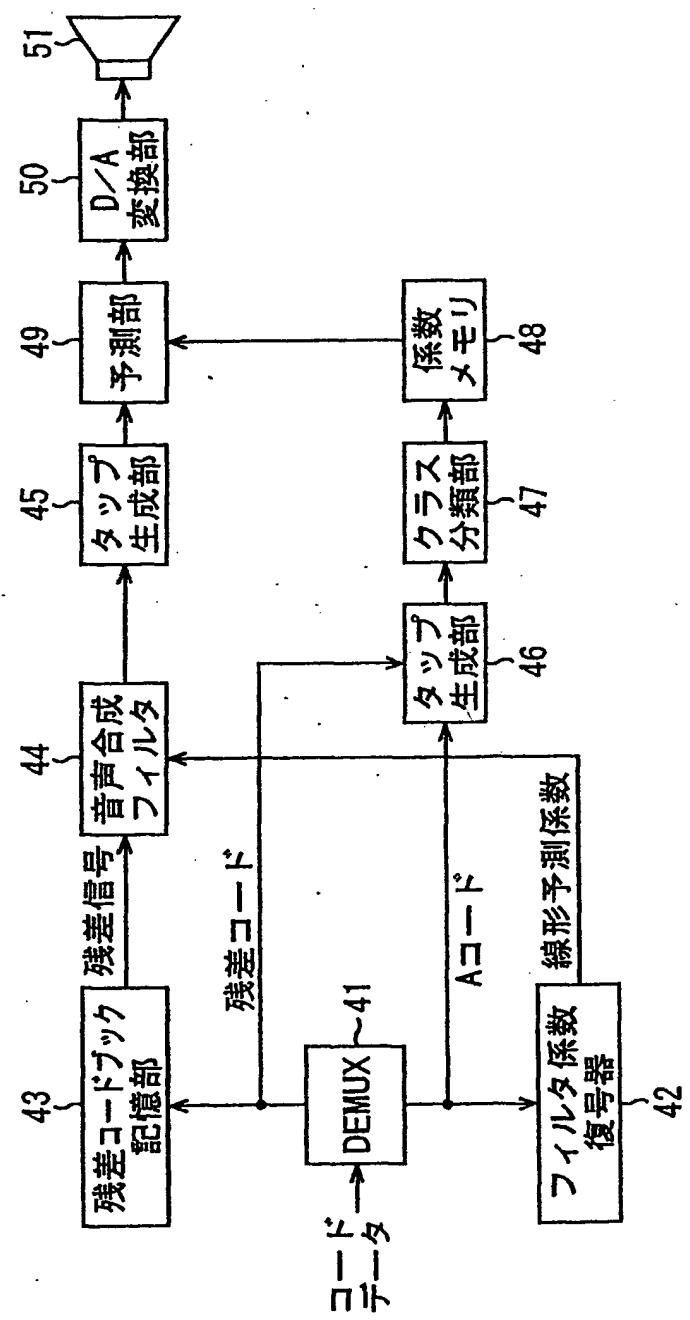


FIG.3

4/34

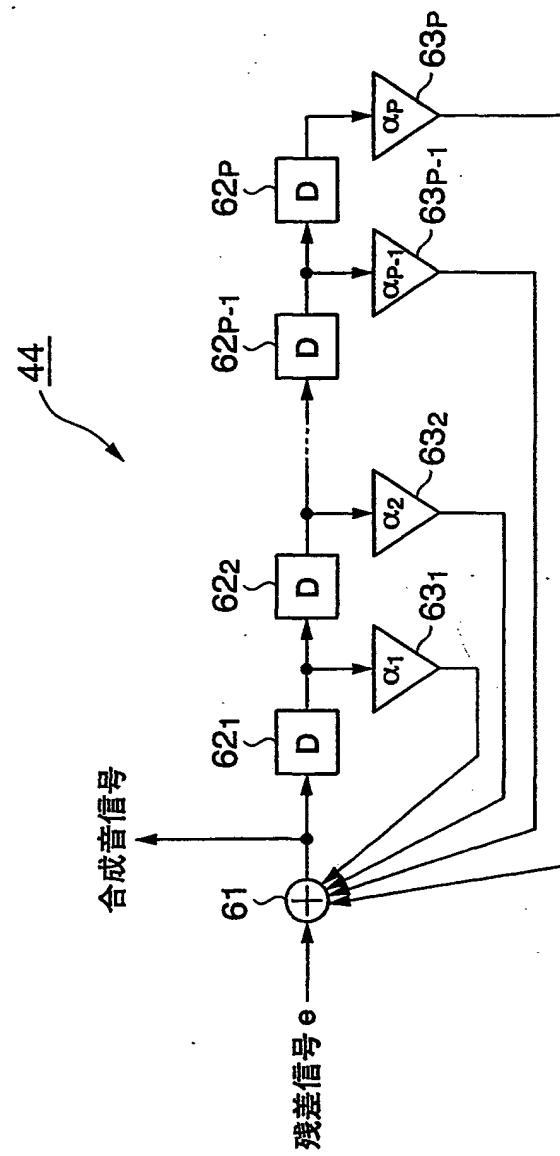


FIG.4

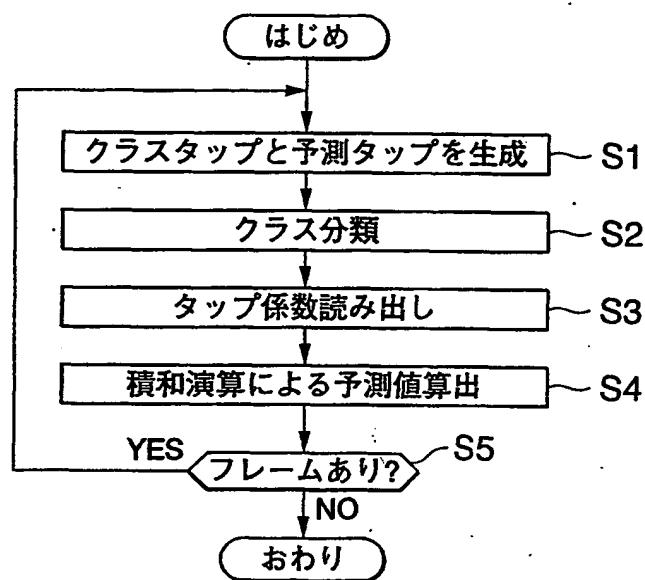


FIG.5

6/34

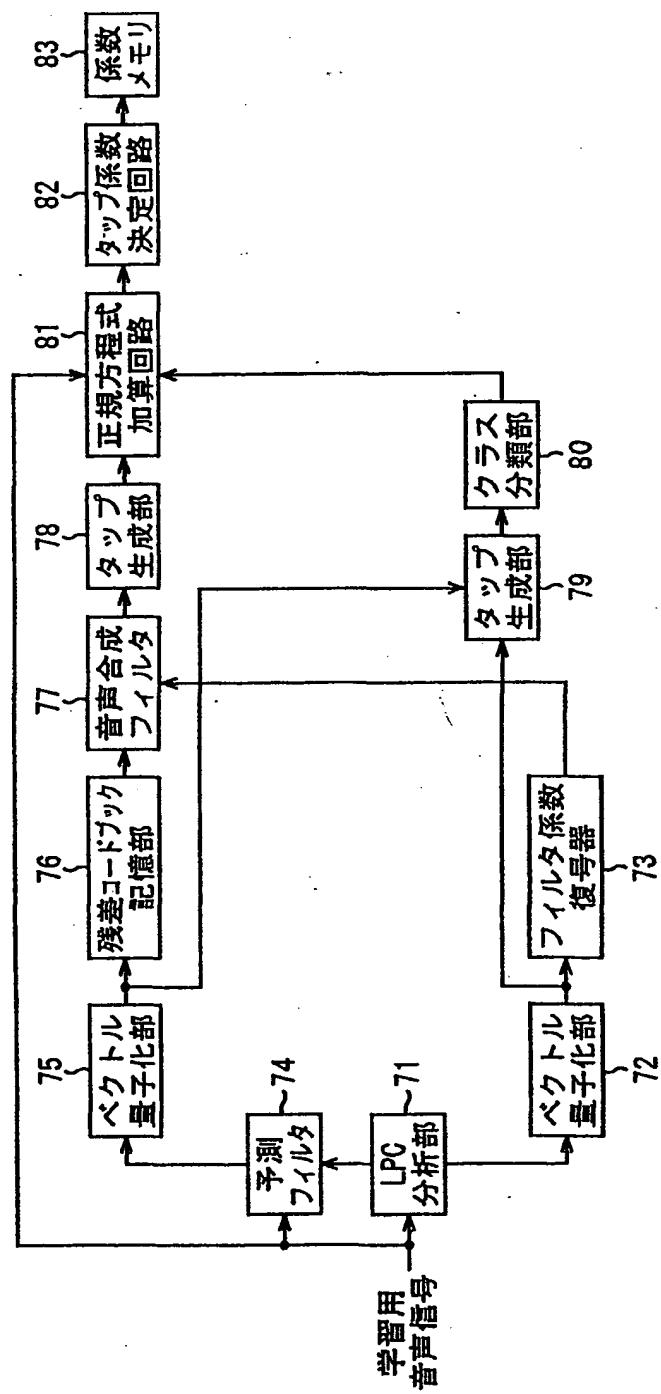


FIG.6

7/34

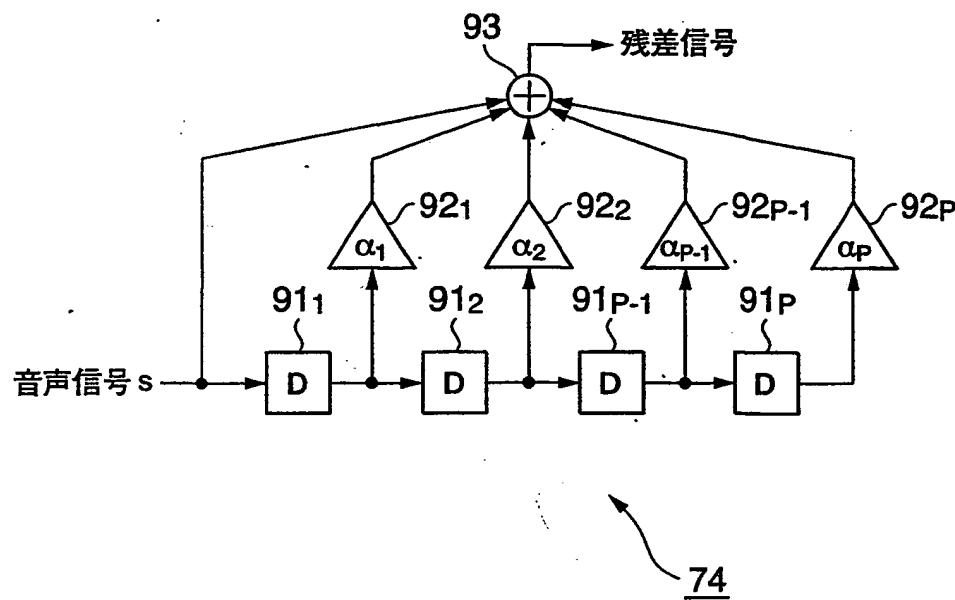


FIG.7

8/34

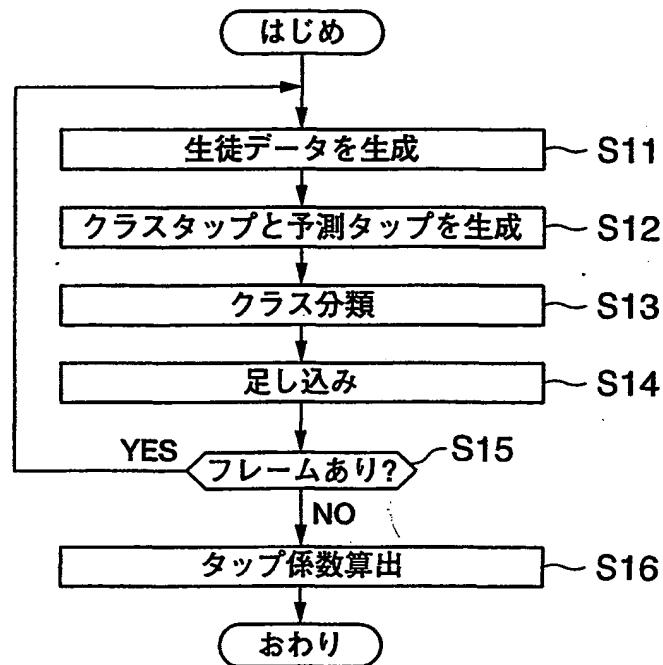


FIG.8

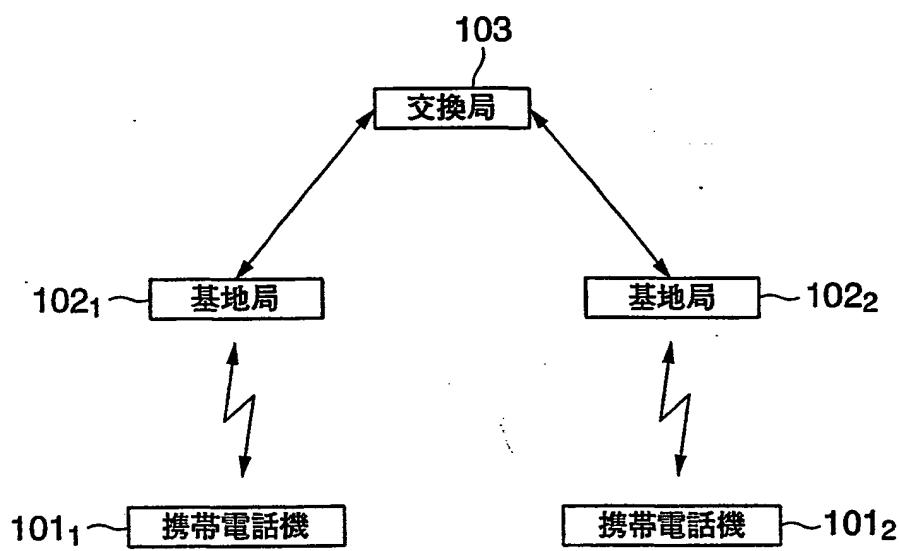


FIG.9

10/34

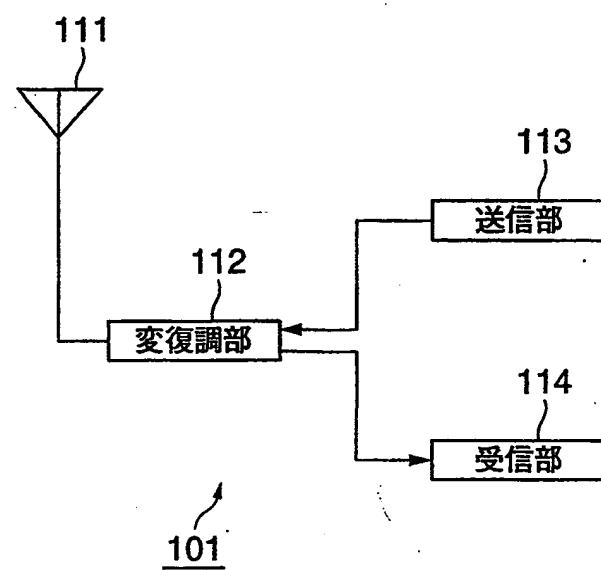


FIG.10

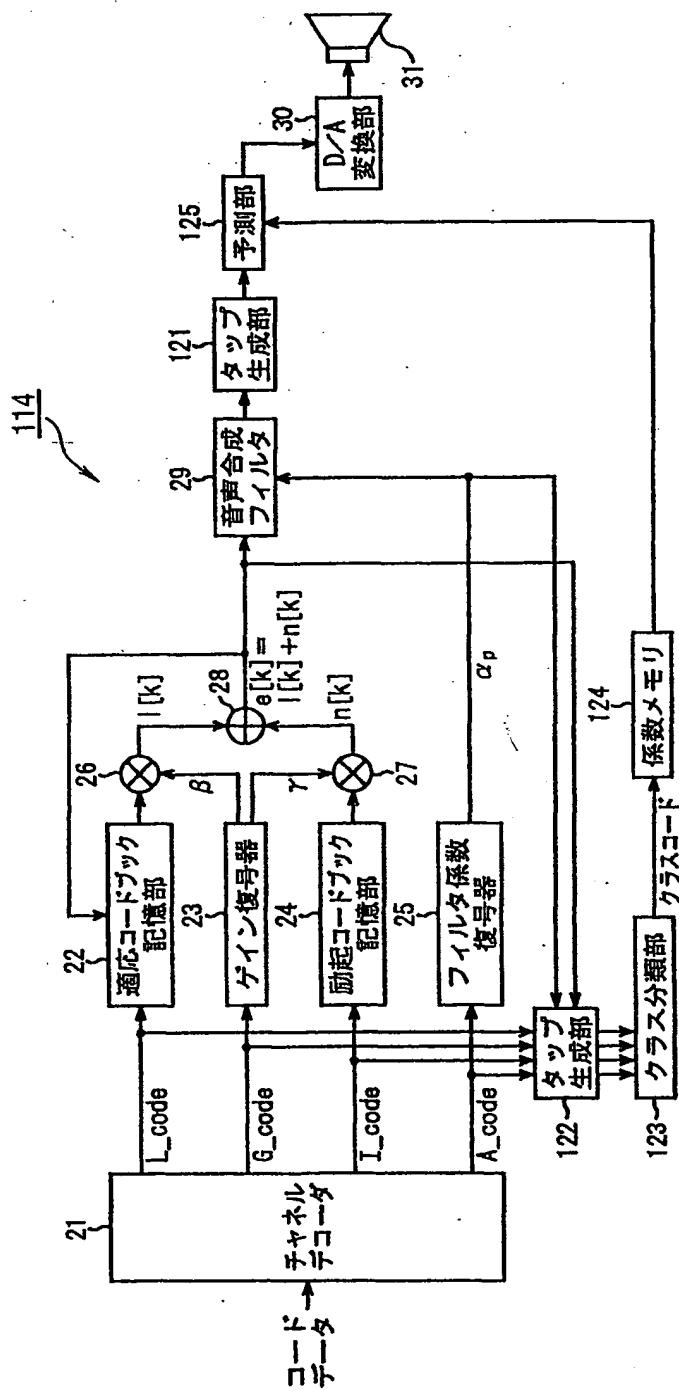


FIG.11

12/34

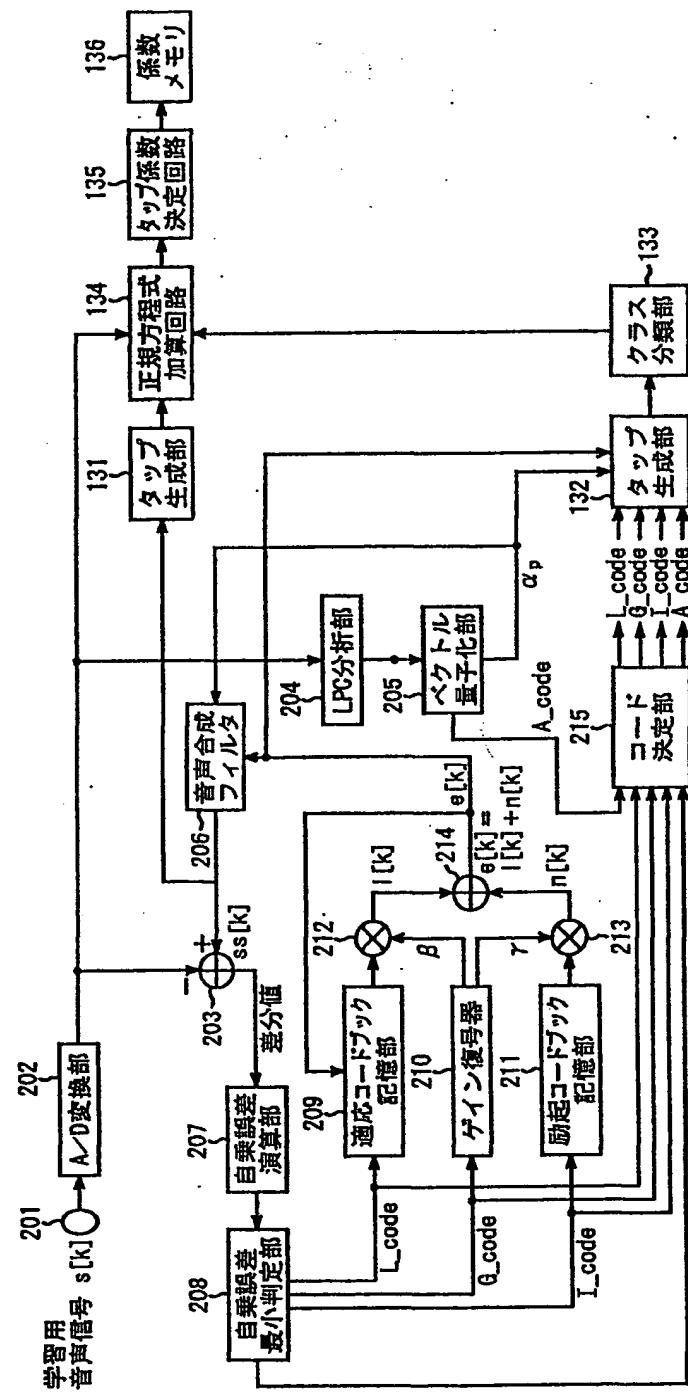


FIG. 12

13/34

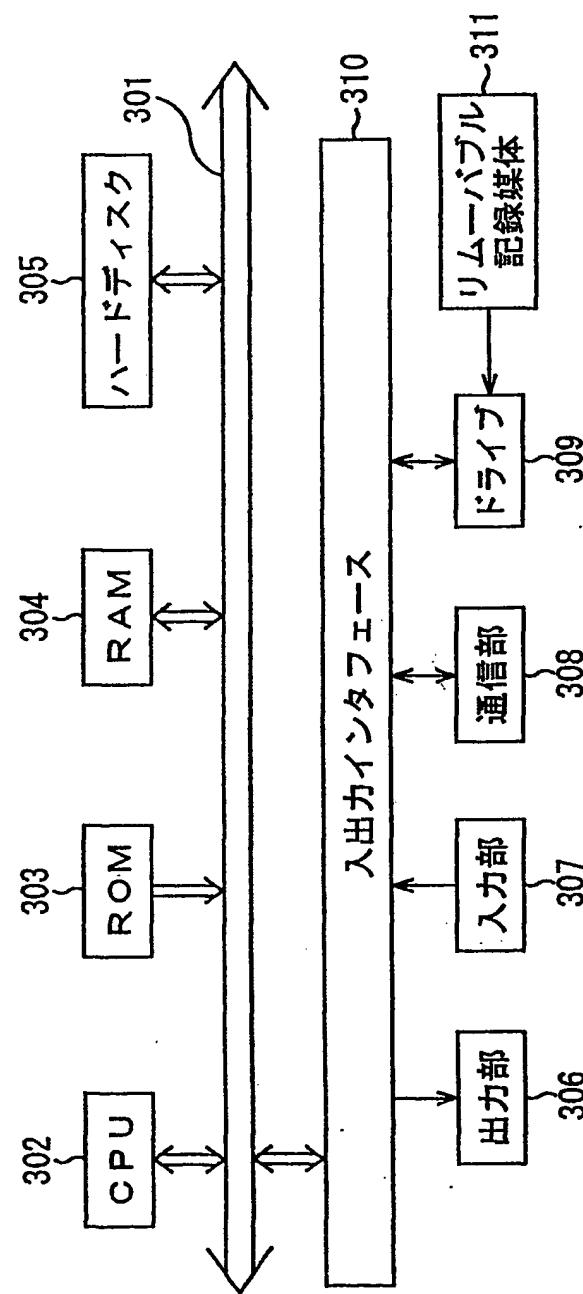


FIG. 13

14/34

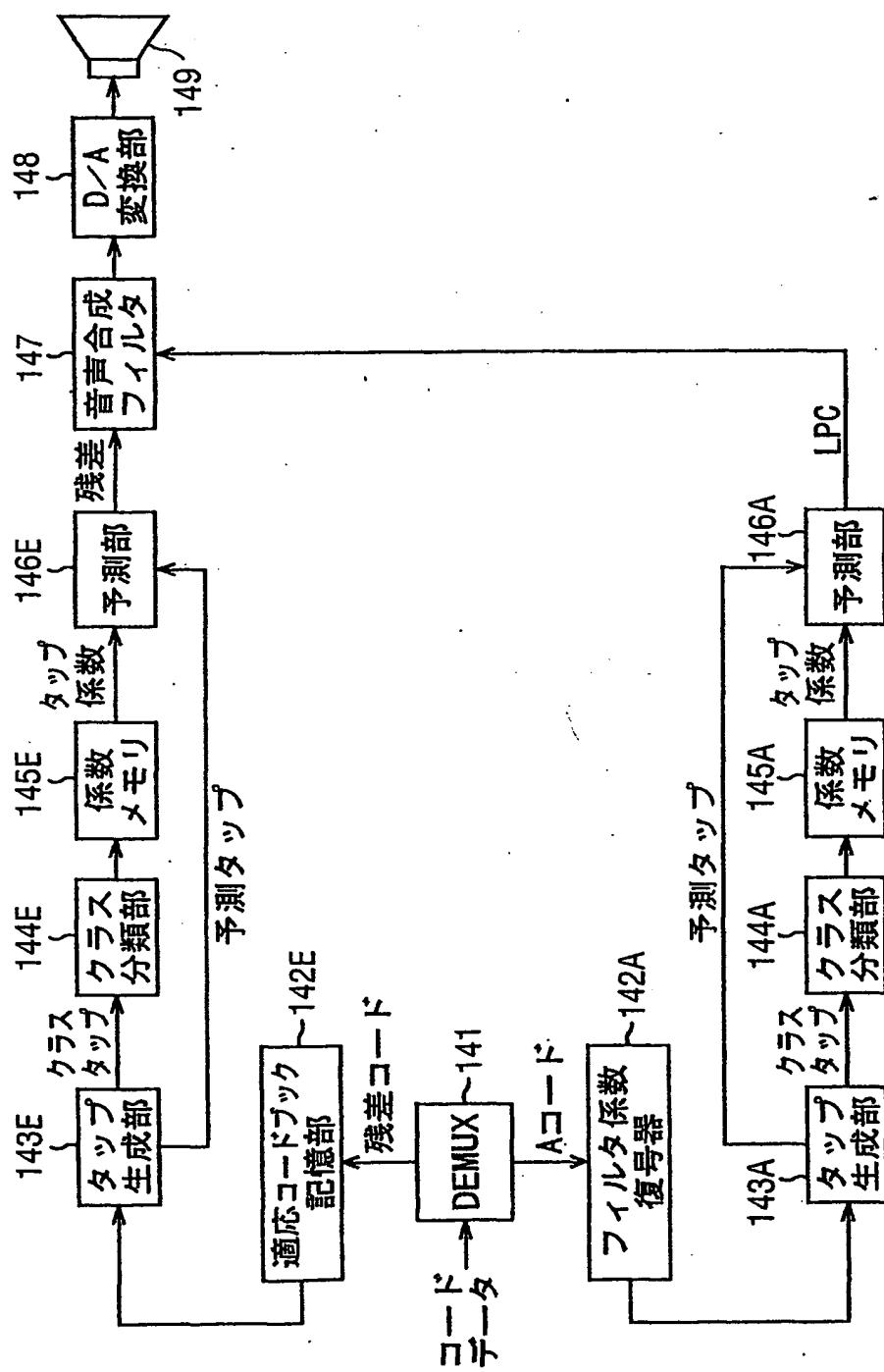


FIG. 14

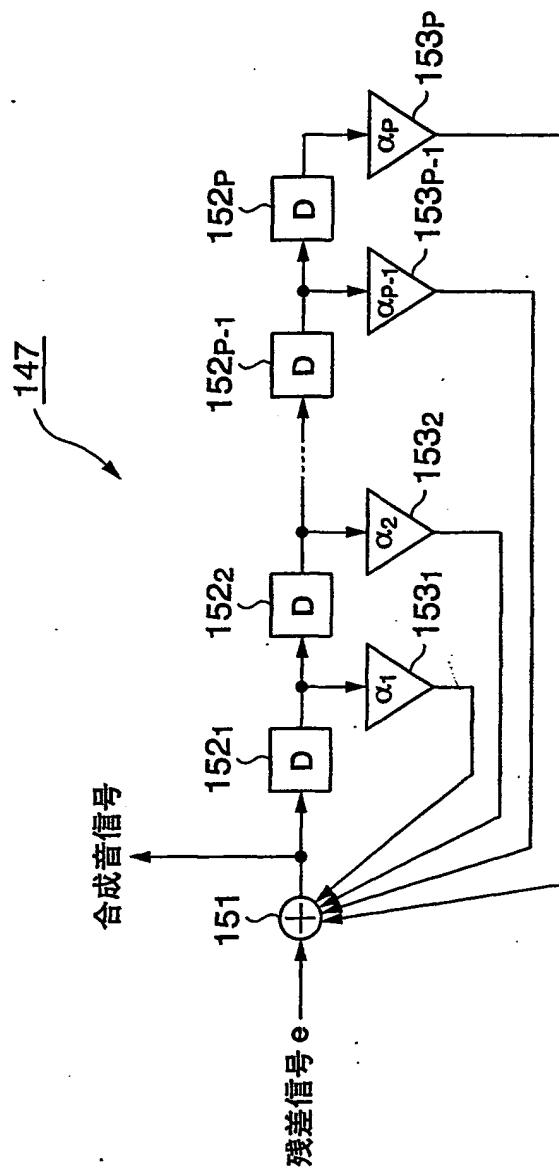


FIG.15

16/34

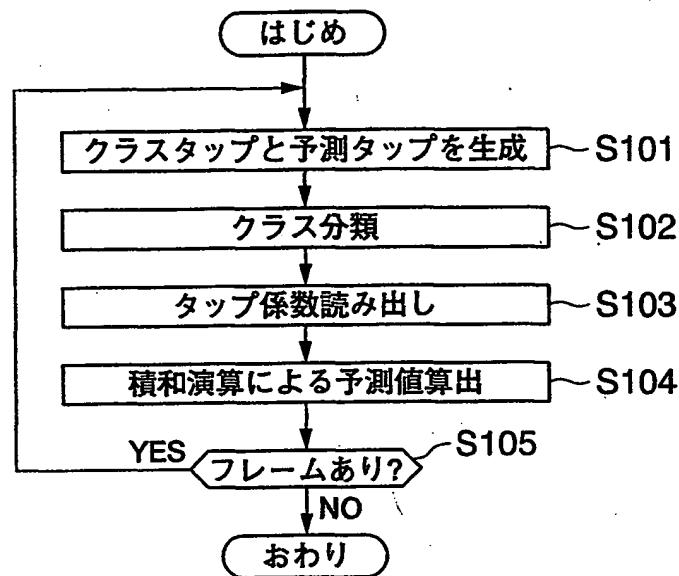


FIG.16

17/34

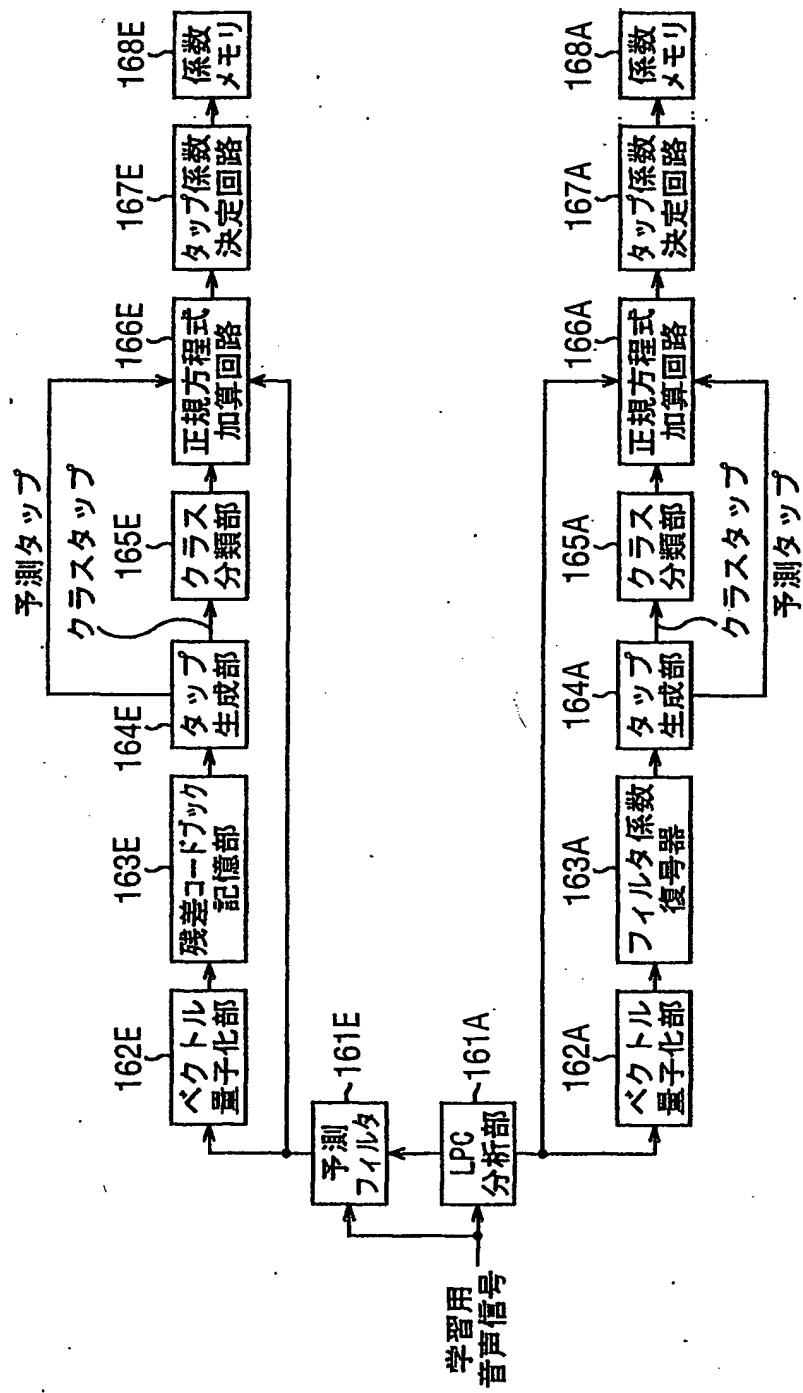


FIG. 17

18/34

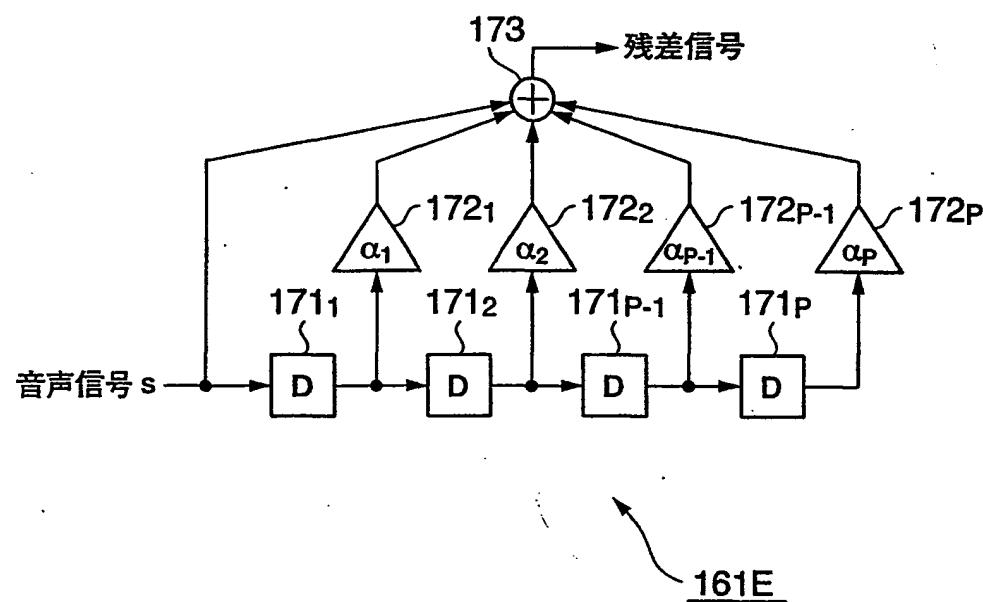


FIG.18

19/34

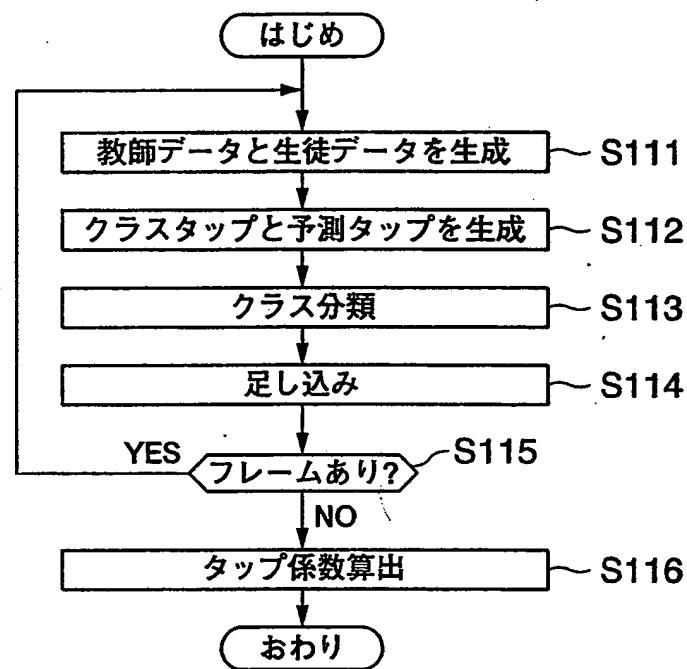


FIG.19

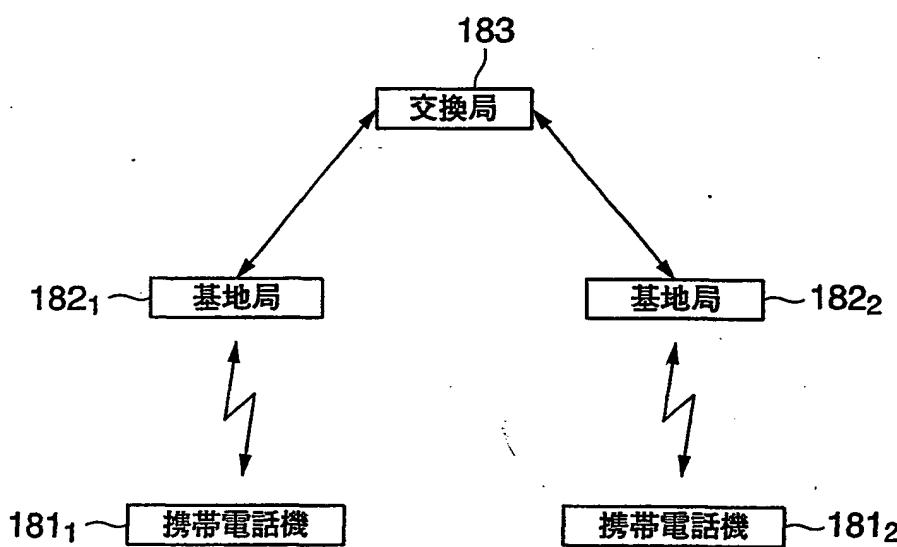


FIG.20

21/34

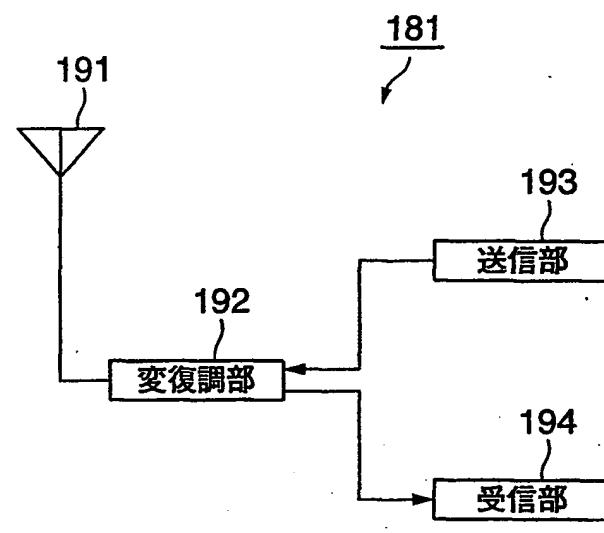


FIG.21

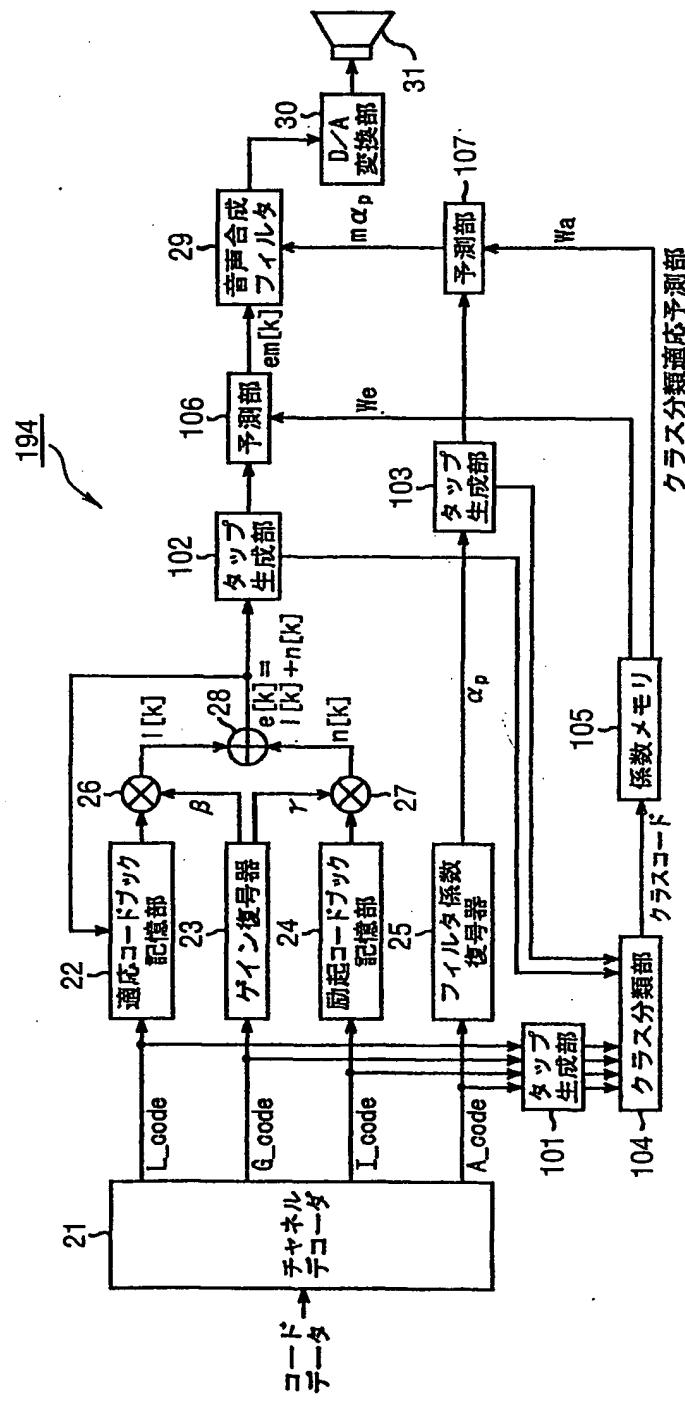


FIG.22

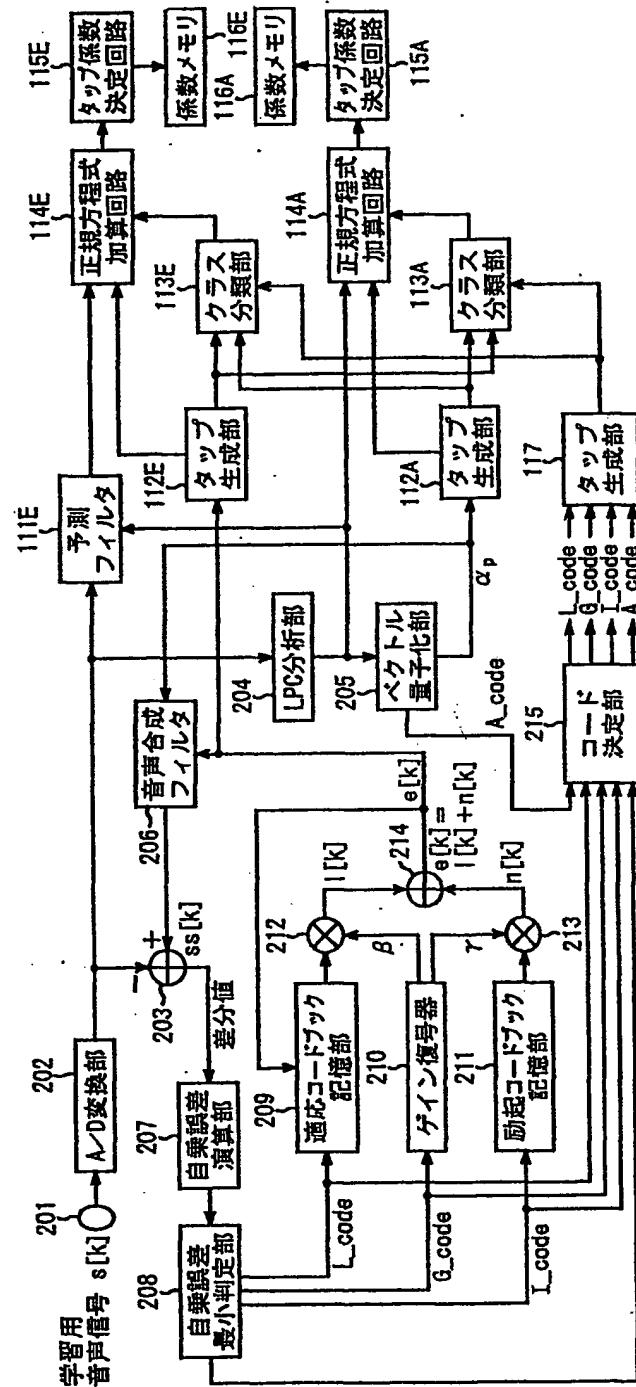


FIG. 23

24/34

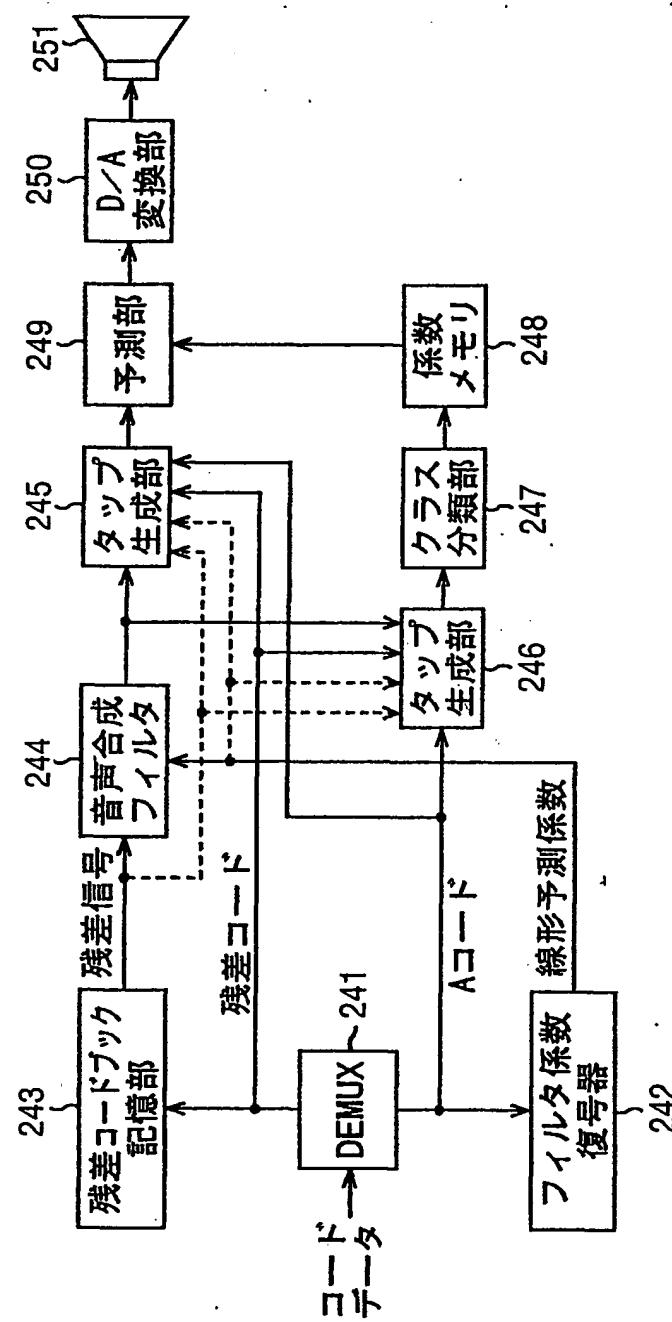


FIG.24

25/34

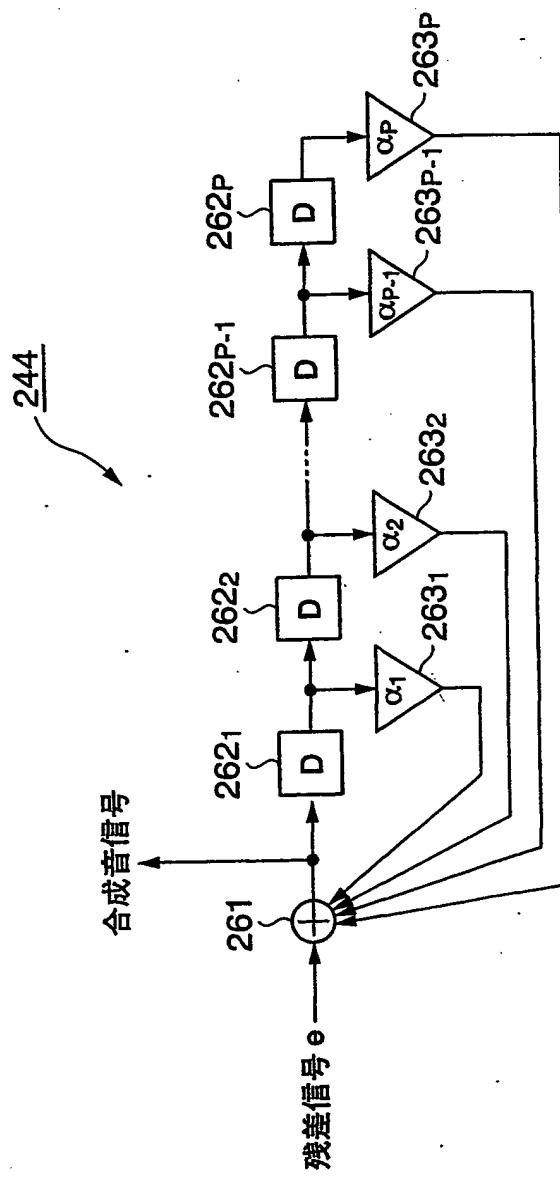


FIG.25

26/34

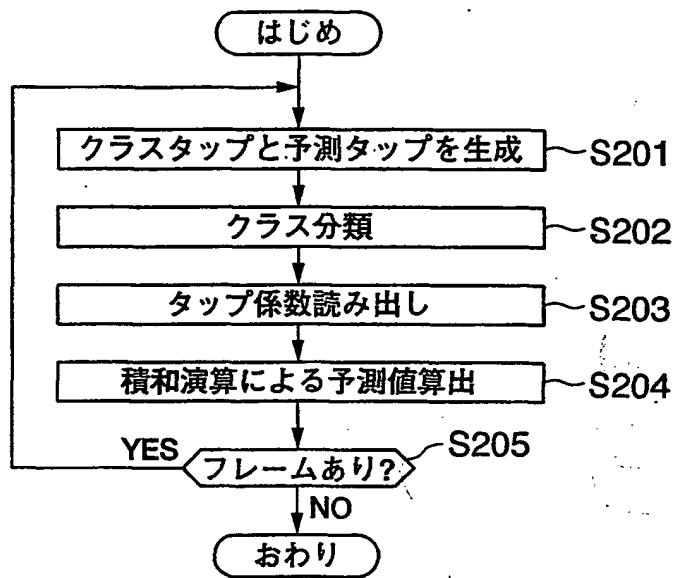


FIG.26

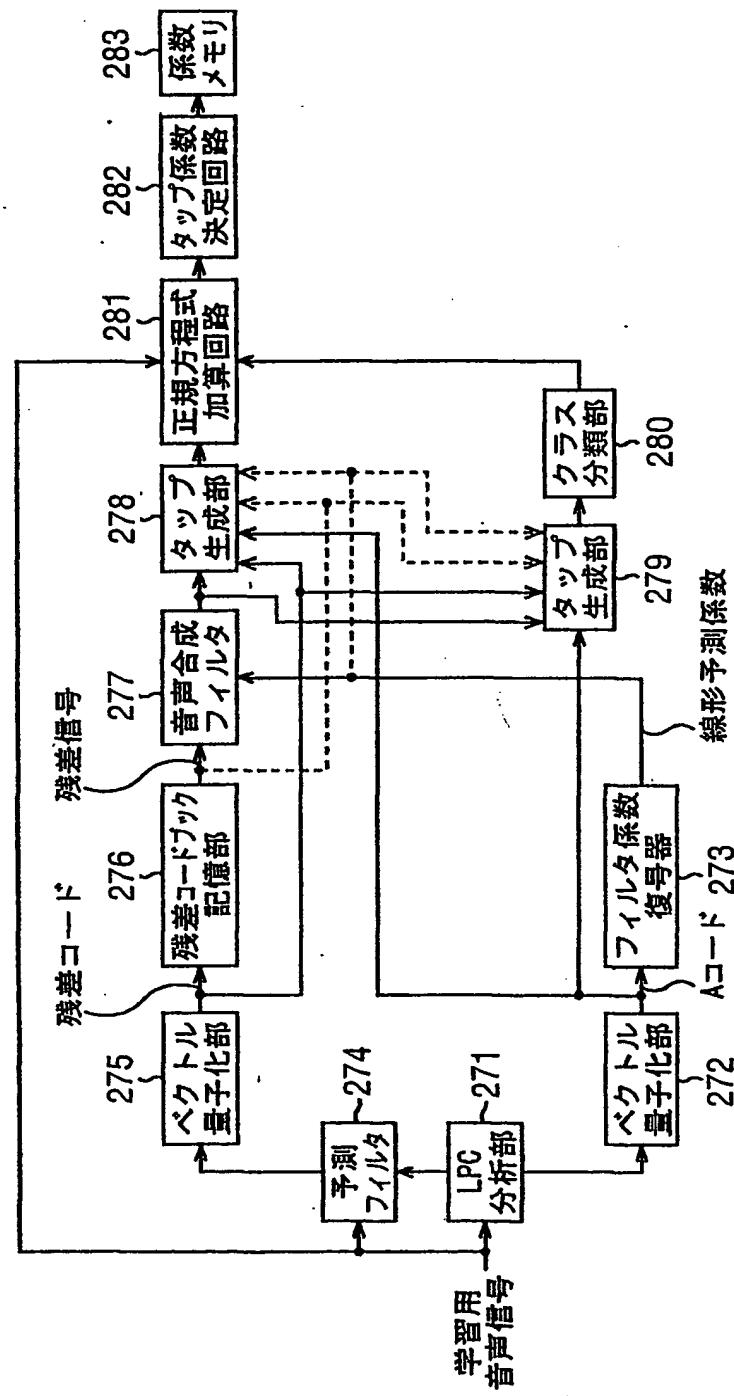


FIG.27

28/34

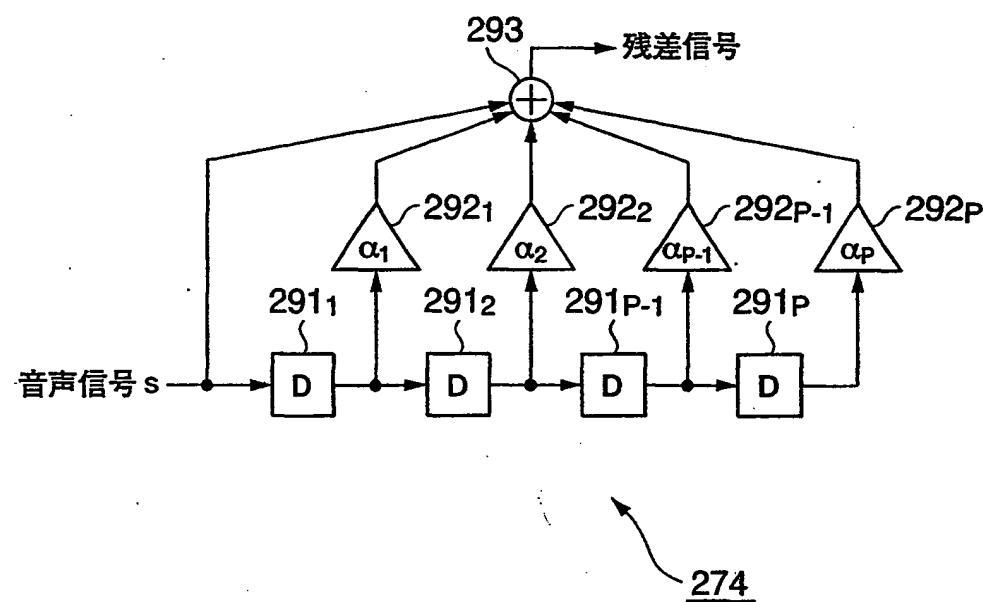


FIG.28

29/34

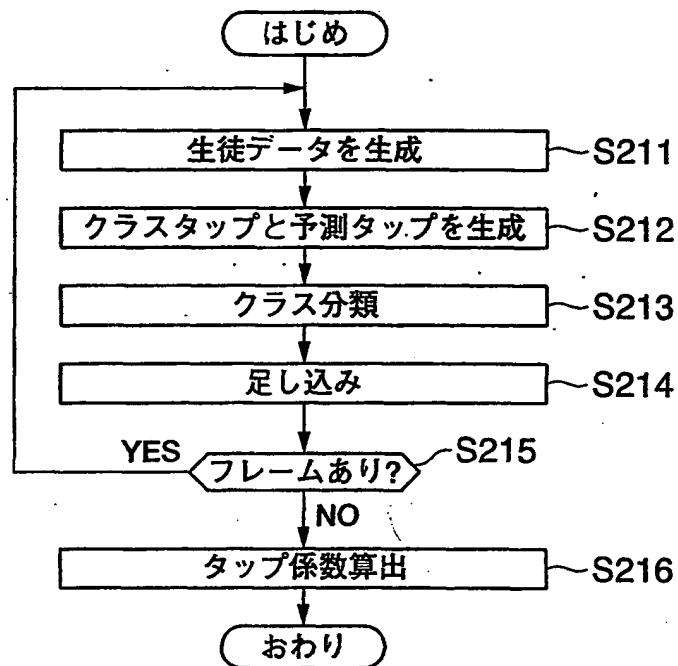


FIG.29

30/34

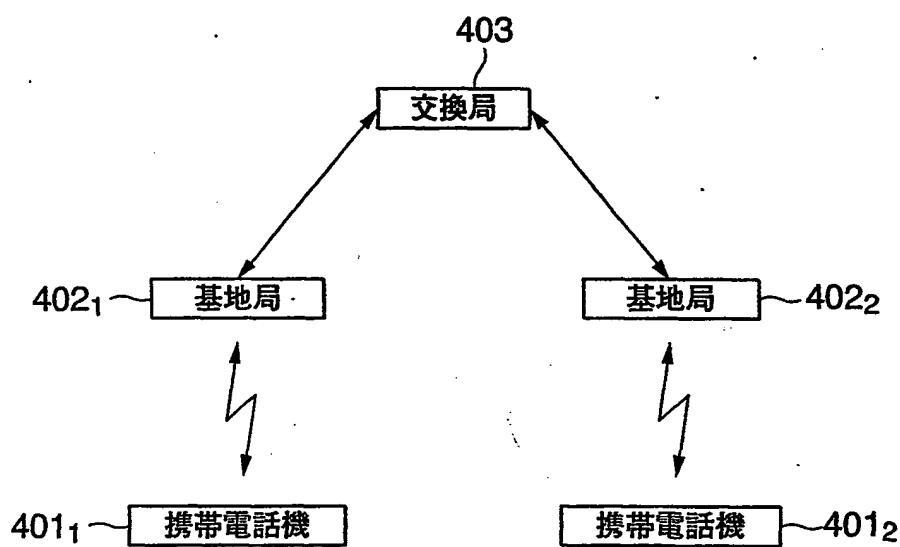


FIG.30

31/34

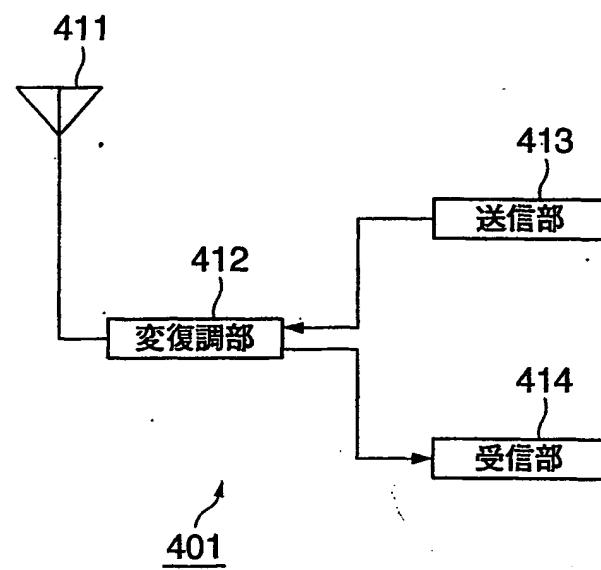


FIG.31

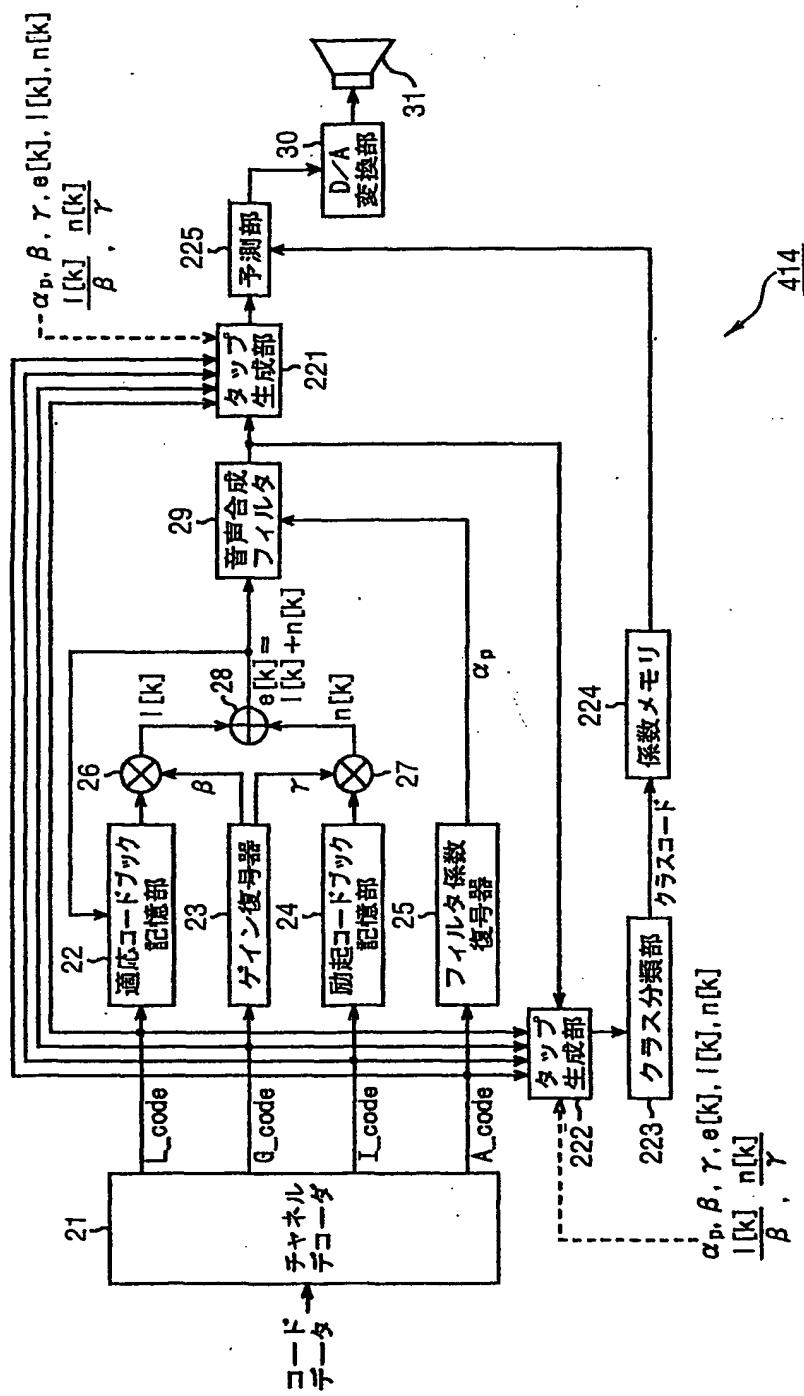


FIG.32

33/34

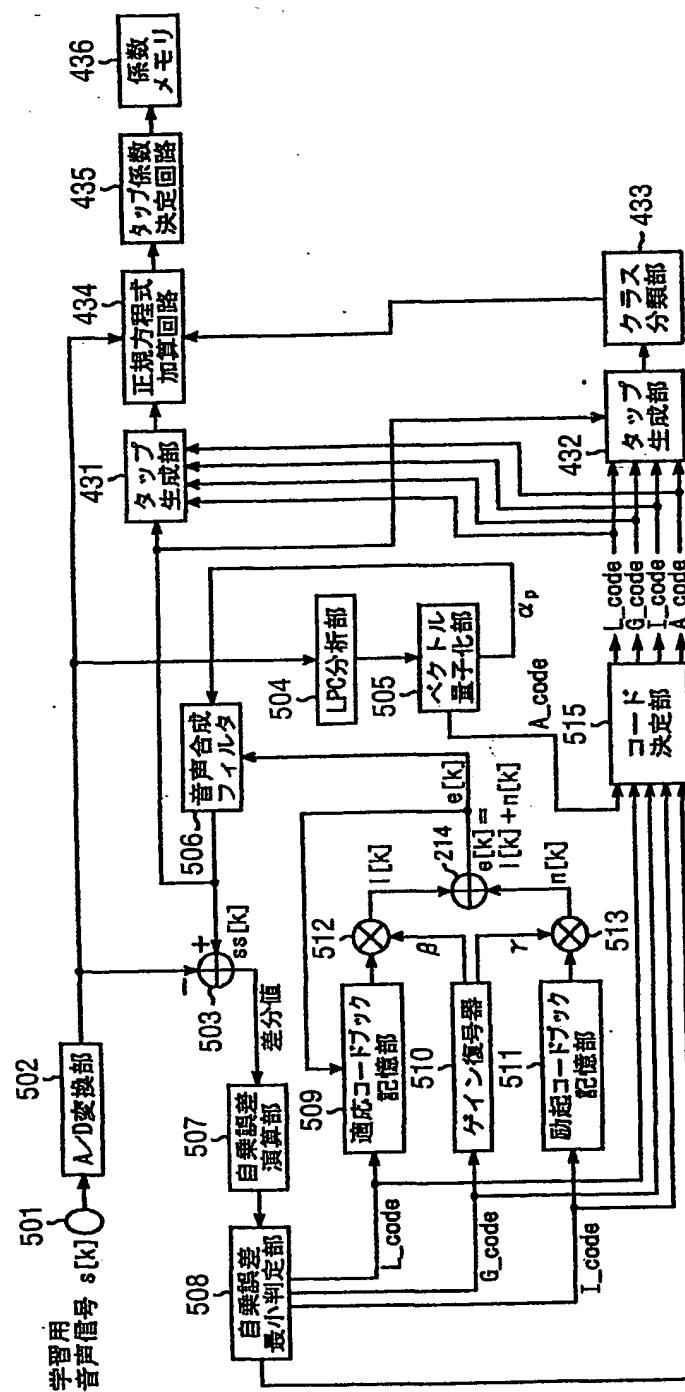


FIG.33

34/34

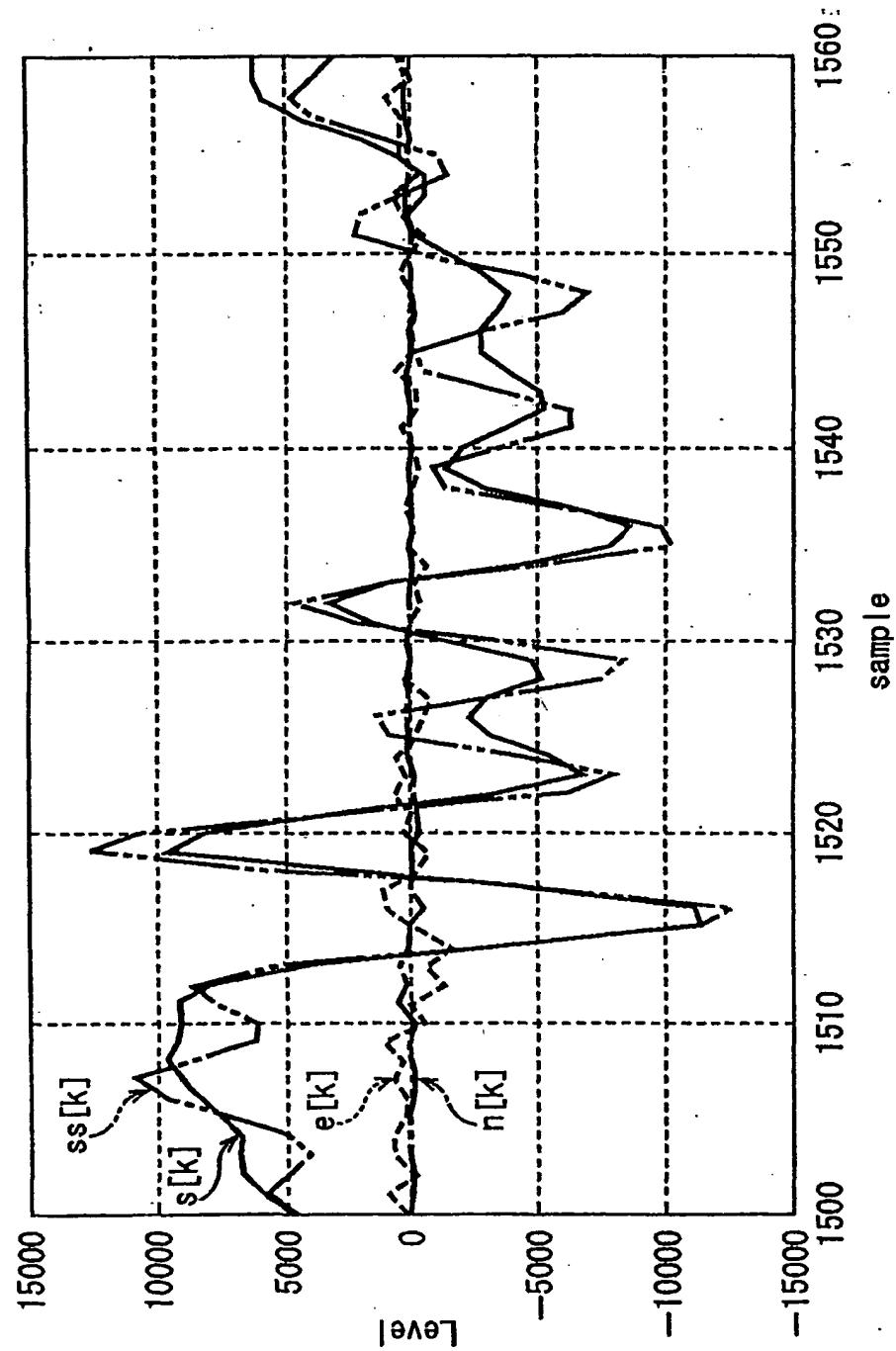


FIG.34

INTERNATIONAL SEARCH REPORT

national application No.

PCT/JP01/06708

A. CLASSIFICATION OF SUBJECT MATTER
 Int.C1' G10L13/00, 19/04 // G10L101:12

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 Int.C1' G10L13/00, 19/00-19/14, H03M7/30, H04B14/04

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
 Jitsuyo Shinan Koho 1922-1996 Toroku Jitsuyo Shinan Koho 1994-2001
 Kokai Jitsuyo Shinan Koho 1971-2001 Jitsuyo Shinan Toroku Koho 1996-2001

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
 JICST FILE (JOIS), WPI (DIALOG)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 10-242867 A (Nippon Telegr. & Teleph. Corp. <NTT>), 11 September, 1998 (11.09.98) (Family: none)	1-53
A	JP 6-83400 A (American Telephone & Telegraph Company), 25 March, 1994 (25.03.94), & US 5327520 A & EP 573216 A2	1-53
A	JP 7-50586 A (AT & T Corporation), 21 February, 1995 (21.02.95), & US 5233660 A & US 5651091 A & US 5680507 A & US 5745871 & EP 532225 A2	1-53
A	JP 8-328591 A (France Telecom), 13 December, 1996 (13.12.96), & FR 2734389 A1 & EP 743634 A1 & US 5845244 A	1-53
A	JP 8-248996 A (Nippon Telegr. & Teleph. Corp. <NTT>), 27 September, 1996 (27.09.96), & EP 731449 A2 & US 5732188 A	1-53

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X"	earlier document but published on or after the international filing date
"E" earlier document but published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
 30 October, 2001 (30.10.01)

Date of mailing of the international search report
 06 November, 2001 (06.11.01)

Name and mailing address of the ISA/
 Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP01/06708

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 9-90997 A (Mitsubishi Electric Corporation), 04 April, 1997 (04.04.97) (Family: none)	1-53
A	JP 9-258795 A (Nippon Telegr. & Teleph. Corp. <NTT>), 03 October, 1997 (03.10.97) (Family: none)	1-53
A	US 6014618 A (DSP Software Engineering, Inc.), 11 January, 2000 (11.01.00) (Family: none)	1-53
A	JP 2000-66700 A (Oki Electric Industry Co., Ltd.), 03 March, 2000 (03.03.00) (Family: none)	1-53

A. 発明の属する分野の分類(国際特許分類(IPC))

Int. Cl' G10L13/00, 19/04
//G10L101:12

B. 調査を行った分野

調査を行った最小限資料(国際特許分類(IPC))

Int. Cl' G10L13/00, 19/00-19/14, H03M7/30, H04B14/04

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年
日本国公開実用新案公報 1971-2001年
日本国登録実用新案公報 1994-2001年
日本国実用新案登録公報 1996-2001年

国際調査で使用した電子データベース(データベースの名称、調査に使用した用語)

JICSTファイル(JOIS), WPI(DIALOG)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 10-242867 A (日本電信電話株式会社), 11. 9 月. 1998 (11. 09. 98) (ファミリーなし)	1-53
A	JP 6-83400 A (アメリカン テレフォン アンド テレグ ラフ カムパニー), 25. 3月. 1994 (25. 03. 94) &US 5327520 A&EP 573216 A2	1-53
A	JP 7-50586 A (エイ・ティ・アンド・ティ・コーポレー ション), 21. 2月. 1995 (21. 02. 95) &US 52336 60 A&US 5651091 A&US 5680507 A&US 5745871 A&EP 532225 A2	1-53

 C欄の続きにも文献が列挙されている。 パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献(理由を付す)

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

30. 10. 01

国際調査報告の発送日

06.11.01

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官(権限のある職員)

山下 岡川史二郎

5C 8946

電話番号 03-3581-1101 内線 3540

C(続き) 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	J P 8-328591 A (フランス・テレコム), 13. 12 月. 1996 (13. 12. 96) &FR 2734389 A1&EP 743634 A1&US 5845244 A	1-53
A	J P 8-248996 A (日本電信電話株式会社), 27. 9 月. 1996 (27. 09. 96) &EP 731449 A2&US 5732188 A	1-53
A	J P 9-90997 A (三菱電機株式会社), 4. 4月. 199 7 (04. 04. 97) (ファミリーなし)	1-53
A	J P 9-258795 A (日本電信電話株式会社), 3. 10 月. 1997 (03. 10. 97) (ファミリーなし)	1-53
A	US 6014618 A (DSP Software Engineering, Inc.), 1 1. 1月. 2000 (11. 01. 00) (ファミリーなし)	1-53
A	J P 2000-66700 A (沖電気工業株式会社), 3. 3 月. 2000 (03. 03. 00) (ファミリーなし)	1-53

国際調査報告

(法8条、法施行規則第40、41条)
(PCT18条、PCT規則43、44)

出願人又は代理人 の書類記号	SK01PCT109	今後の手続きについては、国際調査報告の送付通知様式(PCT/ISA/220)及び下記5を参照すること。	
国際出願番号 PCT/JP01/06708	国際出願日 (日.月.年)	03.08.01	優先日 (日.月.年)
出願人(氏名又は名称) ソニー株式会社			

国際調査機関が作成したこの国際調査報告を法施行規則第41条(PCT18条)の規定に従い出願人に送付する。
この写しは国際事務局にも送付される。

この国際調査報告は、全部で 3 ページである。

この調査報告に引用された先行技術文献の写しも添付されている。

1. 国際調査報告の基礎

a. 言語は、下記に示す場合を除くほか、この国際出願がされたものに基づき国際調査を行った。
 この国際調査機関に提出された国際出願の翻訳文に基づき国際調査を行った。

b. この国際出願は、ヌクレオチド又はアミノ酸配列を含んでおり、次の配列表に基づき国際調査を行った。
 この国際出願に含まれる書面による配列表

この国際出願と共に提出されたフレキシブルディスクによる配列表

出願後に、この国際調査機関に提出された書面による配列表

出願後に、この国際調査機関に提出されたフレキシブルディスクによる配列表

出願後に提出した書面による配列表が出願時における国際出願の開示の範囲を超える事項を含まない旨の陳述書の提出があった。

書面による配列表に記載した配列とフレキシブルディスクによる配列表に記録した配列が同一である旨の陳述書の提出があった。

2. 請求の範囲の一部の調査ができない(第I欄参照)。

3. 発明の單一性が欠如している(第II欄参照)。

4. 発明の名称は 出願人が提出したものと承認する。

次に示すように国際調査機関が作成した。

5. 要約は 出願人が提出したものと承認する。

第III欄に示されているように、法施行規則第47条(PCT規則38.2(b))の規定により国際調査機関が作成した。出願人は、この国際調査報告の発送の日から1カ月以内にこの国際調査機関に意見を提出することができる。

6. 要約書とともに公表される図は、

第 3 図とする。 出願人が示したとおりである。

なし

出願人は図を示さなかった。

本図は発明の特徴を一層よく表している。

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl' G10L13/00, 19/04
//G10L101:12

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl' G10L13/00, 19/00-19/14, H03M7/30, H04B14/04

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年
日本国公開実用新案公報 1971-2001年
日本国登録実用新案公報 1994-2001年
日本国実用新案登録公報 1996-2001年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JICSTファイル (JOIS), WPI (DIALOG)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 10-242867 A (日本電信電話株式会社), 11. 9 月. 1998 (11. 09. 98) (ファミリーなし)	1-53
A	JP 6-83400 A (アメリカン テレフォン アンド テレグラフ カムパニー), 25. 3月. 1994 (25. 03. 94) &US 5327520 A&EP 573216 A2	1-53
A	JP 7-50586 A (エイ・ティ・アンド・ティ・コーポレーション), 21. 2月. 1995 (21. 02. 95) &US 52336 60 A&US 5651091 A&US 5680507 A&US 5745871 A&EP 532225 A2	1-53

C欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

30. 10. 01

国際調査報告の発送日

06.11.01

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

山下 岡里史 (印)

5C 8946

電話番号 03-3581-1101 内線 3540

C (続き) 関連すると認められる文献		関連する 請求の範囲の番号
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	
A	JP 8-328591 A (フランス・テレコム), 13. 12 月. 1996 (13. 12. 96) &FR 2734389 A1&EP 743634 A1&US 5845244 A	1-53
A	JP 8-248996 A (日本電信電話株式会社), 27. 9 月. 1996 (27. 09. 96) &EP 731449 A2&US 5732188 A	1-53
A	JP 9-90997 A (三菱電機株式会社), 4. 4月. 199 7 (04. 04. 97) (ファミリーなし)	1-53
A	JP 9-258795 A (日本電信電話株式会社), 3. 10 月. 1997 (03. 10. 97) (ファミリーなし)	1-53
A	US 6014618 A (DSP Software Engineering, Inc.), 1 1. 1月. 2000 (11. 01. 00) (ファミリーなし)	1-53
A	JP 2000-66700 A (沖電気工業株式会社), 3. 3 月. 2000 (03. 03. 00) (ファミリーなし)	1-53

PATENT COOPERATION TREATY

PCT

NOTIFICATION OF RECEIPT OF
RECORD COPY

(PCT Rule 24.2(a))

From the INTERNATIONAL BUREAU

To:

KOIKE, Akira
 No.11 Mori Bldg., 6-4, Toranomon 2-chome
 Minato-ku, Tokyo 105-0001
 JAPON

Date of mailing (day/month/year) 12 September 2001 (12.09.01)	IMPORTANT NOTIFICATION
Applicant's or agent's file reference SK01PCT109	International application No. PCT/JP01/06708

The applicant is hereby notified that the International Bureau has received the record copy of the international application as detailed below.

Name(s) of the applicant(s) and State(s) for which they are applicants:

SONY CORPORATION (for all designated States except US)
 KONDO, Tetsujiro et al (for US)

International filing date	:	03 August 2001 (03.08.01)
Priority date(s) claimed	:	09 August 2000 (09.08.00)
	:	23 August 2000 (23.08.00)
	:	14 November 2000 (14.11.00)

Date of receipt of the record copy by the International Bureau	:	17 August 2001 (17.08.01)
--	---	---------------------------

List of designated Offices	:	
----------------------------	---	--

EP :AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR
 National :KR,NO,US

ATTENTION

The applicant should carefully check the data appearing in this Notification. In case of any discrepancy between these data and the indications in the international application, the applicant should immediately inform the International Bureau.

In addition, the applicant's attention is drawn to the information contained in the Annex, relating to:

- time limits for entry into the national phase
- confirmation of precautionary designations
- requirements regarding priority documents

A copy of this Notification is being sent to the receiving Office and to the International Searching Authority.

The International Bureau of WIPO
 34, chemin des Colombettes
 1211 Geneva 20, Switzerland

Facsimile No. (41-22) 740.14.35

Authorized officer:



Masashi HONDA

Telephone No. (41-22) 338.83.38

INFORMATION ON TIME LIMITS FOR ENTERING THE NATIONAL PHASE

The applicant is reminded that the "national phase" must be entered before each of the designated Offices indicated in the Notification of Receipt of Record Copy (Form PCT/IB/301) by paying national fees and furnishing translations, as prescribed by the applicable national laws.

The time limit for performing these procedural acts is **20 MONTHS** from the priority date or, for those designated States which the applicant elects in a demand for international preliminary examination or in a later election, **30 MONTHS** from the priority date, provided that the election is made before the expiration of 19 months from the priority date. Some designated (or elected) Offices have fixed time limits which expire even later than 20 or 30 months from the priority date. In other Offices an extension of time or grace period, in some cases upon payment of an additional fee, is available.

In addition to these procedural acts, the applicant may also have to comply with other special requirements applicable in certain Offices. **It is the applicant's responsibility** to ensure that the necessary steps to enter the national phase are taken in a timely fashion. Most designated Offices do not issue reminders to applicants in connection with the entry into the national phase.

For detailed information about the procedural acts to be performed to enter the national phase before each designated Office, the applicable time limits and possible extensions of time or grace periods, and any other requirements, see the relevant Chapters of Volume II of the PCT Applicant's Guide. Information about the requirements for filing a demand for international preliminary examination is set out in Chapter IX of Volume I of the PCT Applicant's Guide.

GR and ES became bound by PCT Chapter II on 7 September 1996 and 6 September 1997, respectively, and may, therefore, be elected in a demand or a later election filed on or after 7 September 1996 and 6 September 1997, respectively, regardless of the filing date of the international application. (See second paragraph above.)

Note that only an applicant who is a national or resident of a PCT Contracting State which is bound by Chapter II has the right to file a demand for international preliminary examination.

CONFIRMATION OF PRECAUTIONARY DESIGNATIONS

This notification lists only specific designations made under Rule 4.9(a) in the request. It is important to check that these designations are correct. Errors in designations can be corrected where precautionary designations have been made under Rule 4.9(b). The applicant is hereby reminded that any precautionary designations may be confirmed according to Rule 4.9(c) before the expiration of 15 months from the priority date. If it is not confirmed, it will automatically be regarded as withdrawn by the applicant. There will be no reminder and no invitation. Confirmation of a designation consists of the filing of a notice specifying the designated State concerned (with an indication of the kind of protection or treatment desired) and the payment of the designation and confirmation fees. Confirmation must reach the receiving Office within the 15-month time limit.

REQUIREMENTS REGARDING PRIORITY DOCUMENTS

For applicants who have not yet complied with the requirements regarding priority documents, the following is recalled.

Where the priority of an earlier national, regional or international application is claimed, the applicant must submit a copy of the said earlier application, certified by the authority with which it was filed ("the priority document") to the receiving Office (which will transmit it to the International Bureau) or directly to the International Bureau, before the expiration of 16 months from the priority date, provided that any such priority document may still be submitted to the International Bureau before that date of international publication of the international application, in which case that document will be considered to have been received by the International Bureau on the last day of the 16-month time limit (Rule 17.1(a)).

Where the priority document is issued by the receiving Office, the applicant may, instead of submitting the priority document, request the receiving Office to prepare and transmit the priority document to the International Bureau. Such request must be made before the expiration of the 16-month time limit and may be subjected by the receiving Office to the payment of a fee (Rule 17.1(b)).

If the priority document concerned is not submitted to the International Bureau or if the request to the receiving Office to prepare and transmit the priority document has not been made (and the corresponding fee, if any, paid) within the applicable time limit indicated under the preceding paragraphs, any designated State may disregard the priority claim, provided that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity to furnish the priority document within a time limit which is reasonable under the circumstances.

Where several priorities are claimed, the priority date to be considered for the purposes of computing the 16-month time limit is the filing date of the earliest application whose priority is claimed.

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

**NOTIFICATION CONCERNING
SUBMISSION OR TRANSMITTAL
OF PRIORITY DOCUMENT**

(PCT Administrative Instructions, Section 411)

Date of mailing (day/month/year) 12 September 2001 (12.09.01)

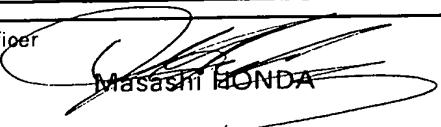
To:

KOIKE, Akira
No.11 Mori Bldg., 6-4, Toranomon 2-chome
Minato-ku, Tokyo 105-0001
JAPON

Applicant's or agent's file reference SK01PCT109	IMPORTANT NOTIFICATION
International application No. PCT/JP01/06708	International filing date (day/month/year) 03 August 2001 (03.08.01)
International publication date (day/month/year) Not yet published	Priority date (day/month/year) 09 August 2000 (09.08.00)
<p>Applicant SONY CORPORATION et al</p>	

1. The applicant is hereby notified of the date of receipt (except where the letters "NR" appear in the right-hand column) by the International Bureau of the priority document(s) relating to the earlier application(s) indicated below. Unless otherwise indicated by an asterisk appearing next to a date of receipt, or by the letters "NR", in the right-hand column, the priority document concerned was submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b).
2. This updates and replaces any previously issued notification concerning submission or transmittal of priority documents.
3. An asterisk(*) appearing next to a date of receipt, in the right-hand column, denotes a priority document submitted or transmitted to the International Bureau but not in compliance with Rule 17.1(a) or (b). In such a case, the attention of the applicant is directed to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.
4. The letters "NR" appearing in the right-hand column denote a priority document which was not received by the International Bureau or which the applicant did not request the receiving Office to prepare and transmit to the International Bureau, as provided by Rule 17.1(a) or (b), respectively. In such a case, the attention of the applicant is directed to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.

<u>Priority date</u>	<u>Priority application No.</u>	<u>Country or regional Office or PCT receiving Office</u>	<u>Date of receipt of priority document</u>
09 Augu 2000 (09.08.00)	2000-241062	JP	17 Augu 2001 (17.08.01)

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No. (41-22) 740.14.35	Authorized officer  Masashi HONDA Telephone No. (41-22) 338.83.38
--	--

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

**NOTIFICATION CONCERNING
SUBMISSION OR TRANSMITTAL
OF PRIORITY DOCUMENT**

(PCT Administrative Instructions, Section 411)

Date of mailing (day/month/year) 19 October 2001 (19.10.01)
Applicant's or agent's file reference SK01PCT109
International application No. PCT/JP01/06708
International publication date (day/month/year) Not yet published
Applicant SONY CORPORATION et al

To:

KOIKE, Akira
No.11 Mori Bldg., 6-4, Toranomon 2-chome
Minato-ku, Tokyo 105-0001
JAPON

IMPORTANT NOTIFICATION

International filing date (day/month/year)
03 August 2001 (03.08.01)

Priority date (day/month/year)
09 August 2000 (09.08.00)

- The applicant is hereby notified of the date of receipt (except where the letters "NR" appear in the right-hand column) by the International Bureau of the priority document(s) relating to the earlier application(s) indicated below. Unless otherwise indicated by an asterisk appearing next to a date of receipt, or by the letters "NR", in the right-hand column, the priority document concerned was submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b).
- This updates and replaces any previously issued notification concerning submission or transmittal of priority documents.
- An asterisk(*) appearing next to a date of receipt, in the right-hand column, denotes a priority document submitted or transmitted to the International Bureau but not in compliance with Rule 17.1(a) or (b). In such a case, **the attention of the applicant is directed** to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.
- The letters "NR" appearing in the right-hand column denote a priority document which was not received by the International Bureau or which the applicant did not request the receiving Office to prepare and transmit to the International Bureau, as provided by Rule 17.1(a) or (b), respectively. In such a case, **the attention of the applicant is directed** to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.

<u>Priority date</u>	<u>Priority application No.</u>	<u>Country or regional Office or PCT receiving Office</u>	<u>Date of receipt of priority document</u>
09 Augu 2000 (09.08.00)	2000-241062	JP	17 Augu 2001 (17.08.01)
23 Augu 2000 (23.08.00)	2000-251969	JP	21 Sept 2001 (21.09.01)
14 Nove 2000 (14.11.00)	2000-346675	JP	21 Sept 2001 (21.09.01)

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No. (41-22) 740.14.35	Authorized officer Khemais BRAHMI Telephone No. (41-22) 338.83.38
--	---

004385132

PATENT COOPERATION TREATY

PCT

NOTICE INFORMING THE APPLICANT OF THE
COMMUNICATION OF THE INTERNATIONAL
APPLICATION TO THE DESIGNATED OFFICES

(PCT Rule 47.1(c), first sentence)

From the INTERNATIONAL BUREAU

To:

KOIKE, Akira
No.11 Mori Bldg., 6-4, Toranomon 2-chome
Minato-ku, Tokyo 105-0001
JAPON

Date of mailing (day/month/year) 14 February 2002 (14.02.02)
--

Applicant's or agent's file reference SK01PCT109
--

IMPORTANT NOTICE

International application No. PCT/JP01/06708	International filing date (day/month/year) 03 August 2001 (03.08.01)	Priority date (day/month/year) 09 August 2000 (09.08.00)
Applicant SONY CORPORATION et al		

1. Notice is hereby given that the International Bureau has **communicated**, as provided in Article 20, the international application to the following designated Offices on the date indicated above as the date of mailing of this notice:
KR,US

In accordance with Rule 47.1(c), third sentence, those Offices will accept the present notice as conclusive evidence that the communication of the international application has duly taken place on the date of mailing indicated above and no copy of the international application is required to be furnished by the applicant to the designated Office(s).

2. The following designated Offices have waived the requirement for such a communication at this time:
EP,NO

The communication will be made to those Offices only upon their request. Furthermore, those Offices do not require the applicant to furnish a copy of the international application (Rule 49.1(a-bis)).

3. Enclosed with this notice is a copy of the international application as published by the International Bureau on
14 February 2002 (14.02.02) under No. WO 02/13183

REMINDER REGARDING CHAPTER II (Article 31(2)(a) and Rule 54.2)

If the applicant wishes to postpone entry into the national phase until 30 months (or later in some Offices) from the priority date, a **demand for international preliminary examination** must be filed with the competent International Preliminary Examining Authority before the expiration of 19 months from the priority date.

It is the applicant's sole responsibility to monitor the 19-month time limit.

Note that only an applicant who is a national or resident of a PCT Contracting State which is bound by Chapter II has the right to file a demand for international preliminary examination (at present, all PCT Contracting States are bound by Chapter II).

REMINDER REGARDING ENTRY INTO THE NATIONAL PHASE (Article 22 or 39(1))

If the applicant wishes to proceed with the international application in the **national phase**, he must, within 20 months or 30 months, or later in some Offices, perform the acts referred to therein before each designated or elected Office.

For further important information on the time limits and acts to be performed for entering the national phase, see the Annex to Form PCT/IB/301 (Notification of Receipt of Record Copy) and the PCT Applicant's Guide, Volume II.

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No. (41-22) 740.14.35

Authorized officer

J. Zahra

Telephone No. (41-22) 338.91.11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP01/06708

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl' G10L13/00, 19/04 // G10L101:12

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl' G10L13/00, 19/00-19/14, H03M7/30, H04B14/04

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2001
Kokai Jitsuyo Shinan Koho	1971-2001	Jitsuyo Shinan Toroku Koho	1996-2001

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JICST FILE (JOIS), WPI (DIALOG)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 10-242867 A (Nippon Telegr. & Teleph. Corp. <NTT>), 11 September, 1998 (11.09.98) (Family: none)	1-53
A	JP 6-83400 A (American Telephone & Telegraph Company), 25 March, 1994 (25.03.94), & US 5327520 A & EP 573216 A2	1-53
A	JP 7-50586 A (AT & T Corporation), 21 February, 1995 (21.02.95), & US 5233660 A & US 5651091 A & US 5680507 A & US 5745871 & EP 532225 A2	1-53
A	JP 8-328591 A (France Telecom), 13 December, 1996 (13.12.96), & FR 2734389 A1 & EP 743634 A1 & US 5845244 A	1-53
A	JP 8-248996 A (Nippon Telegr. & Teleph. Corp. <NTT>), 27 September, 1996 (27.09.96), & EP 731449 A2 & US 5732188 A	1-53

 Further documents are listed in the continuation of Box C. See patent family annex.

* "A" Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier document but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
30 October, 2001 (30.10.01)Date of mailing of the international search report
06 November, 2001 (06.11.01)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP01/06708

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 9-90997 A (Mitsubishi Electric Corporation), 04 April, 1997 (04.04.97) (Family: none)	1-53
A	JP 9-258795 A (Nippon Telegr. & Teleph. Corp. <NTT>), 03 October, 1997 (03.10.97) (Family: none)	1-53
A	US 6014618 A (DSP Software Engineering, Inc.), 11 January, 2000 (11.01.00) (Family: none)	1-53
A	JP 2000-66700 A (Oki Electric Industry Co., Ltd.), 03 March, 2000 (03.03.00) (Family: none)	1-53